

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

## Journal of Differential Equations

[www.elsevier.com/locate/jde](http://www.elsevier.com/locate/jde)

## Ghost circles in lattice Aubry–Mather theory

Blaz Mramor\*, Bob Rink

Department of Mathematics, VU University Amsterdam, The Netherlands

## ARTICLE INFO

## Article history:

Received 11 January 2011

Revised 20 November 2011

Available online 10 December 2011

## ABSTRACT

Monotone lattice recurrence relations such as the Frenkel–Kontorova lattice, arise in Hamiltonian lattice mechanics, as models for ferromagnetism and as discretization of elliptic PDEs. Mathematically, they are a multi-dimensional counterpart of monotone twist maps.

Such recurrence relations often admit a variational structure, so that the solutions  $x : \mathbb{Z}^d \rightarrow \mathbb{R}$  are the stationary points of a formal action function  $W(x)$ . Given any rotation vector  $\omega \in \mathbb{R}^d$ , classical Aubry–Mather theory establishes the existence of a large collection of solutions of  $\nabla W(x) = 0$  of rotation vector  $\omega$ . For irrational  $\omega$ , this is the well-known Aubry–Mather set. It consists of global minimizers and it may have gaps.

In this paper, we study the parabolic gradient flow  $\frac{dx}{dt} = -\nabla W(x)$  and we will prove that every Aubry–Mather set can be interpolated by a continuous gradient-flow invariant family, the so-called ‘ghost circle’. The existence of these ghost circles is known in dimension  $d = 1$ , for rational rotation vectors and Morse action functions. The main technical result of this paper is therefore a compactness theorem for lattice ghost circles, based on a parabolic Harnack inequality for the gradient flow. This implies the existence of lattice ghost circles of arbitrary rotation vectors and for arbitrary actions.

As a consequence, we can give a simple proof of the fact that when an Aubry–Mather set has a gap, then this gap must be filled with minimizers, or contain a non-minimizing solution.

© 2011 Elsevier Inc. All rights reserved.

\* Corresponding author.

E-mail addresses: [bmramor@few.vu.nl](mailto:bmramor@few.vu.nl) (B. Mramor), [brink@few.vu.nl](mailto:brink@few.vu.nl) (B. Rink).

## 1. Introduction and outline

In this paper we are interested in variational monotone lattice recurrence relations. Before introducing such recurrence relations in full generality, let us discuss as an example the so-called  $d$ -dimensional Frenkel–Kontorova lattice. Here, the goal is to find a  $d$ -dimensional “lattice configuration”  $x: \mathbb{Z}^d \rightarrow \mathbb{R}$  that satisfies

$$V'(x_i) - (\Delta x)_i = 0 \quad \text{for all } i \in \mathbb{Z}^d. \quad (1.1)$$

In the equation above, the smooth function  $V: \mathbb{R} \rightarrow \mathbb{R}$  satisfies  $V(\xi + 1) = V(\xi)$  for all  $\xi \in \mathbb{R}$ . It has the interpretation of a periodic onsite potential. Setting  $\|i\| := \sum_{k=1}^d |i_k|$ , the discrete Laplace operator  $\Delta: \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$  is defined as

$$(\Delta x)_i := \frac{1}{2d} \sum_{\|j-i\|=1} (x_j - x_i) \quad \text{for all } i \in \mathbb{Z}^d. \quad (1.2)$$

One could think of Eq. (1.1) as a naive discretization of the nonlinear elliptic partial differential equation  $V'(u) - \Delta u = 0$  for a function  $u: \mathbb{R}^d \rightarrow \mathbb{R}$  and  $x_i = u(i)$ .

At the same time, Eq. (1.1) is relevant for statistical mechanics, because it is related to the Frenkel–Kontorova Hamiltonian lattice differential equation

$$\frac{d^2 x_i}{dt^2} + V'(x_i) - (\Delta x)_i = 0 \quad \text{for all } i \in \mathbb{Z}^d. \quad (1.3)$$

This differential equation describes the motion of particles under the competing influence of an onsite periodic potential field and nearest neighbor attraction. Obviously, Eq. (1.1) describes its stationary solutions.

Finally, in dimension  $d = 1$ , the solutions of Eq. (1.1) correspond to orbits of the famous Chirikov standard map  $T_V$  of the annulus. This correspondence is explained in some detail in Appendix A.

The Frenkel–Kontorova problem (1.1) is an example from a quite general class of lattice recurrence relations to which the results of this paper apply. These are recurrence relations for which there exists, for every  $j \in \mathbb{Z}^d$ , a real-valued “local potential” function  $S_j: \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  so that the relation can be written in the form

$$\sum_{j \in \mathbb{Z}^d} \partial_i S_j(x) = 0 \quad \text{for all } i \in \mathbb{Z}^d. \quad (1.4)$$

It turns out that for the Frenkel–Kontorova problem (1.1), such local potentials exist and it is easy to check that they are given by

$$S_j(x) := V(x_j) + \frac{1}{8d} \sum_{\|k-j\|=1} (x_k - x_j)^2. \quad (1.5)$$

For the general problem (1.4), the functions  $S_j(x)$  will be required to satisfy some rather restrictive hypotheses that will be explained in detail in Section 2. Physically, the most important of these hypotheses is the *monotonicity* condition. It is a discrete analogue of ellipticity for a PDE. Among the more technical hypotheses is one that guarantees that the sums in expression (1.4) are finite. For the purpose of this introduction, it probably suffices to say that the potentials (1.5) of Frenkel–Kontorova are prototypical for the  $S_j(x)$  that we have in mind.

It is important to observe that the solutions of (1.4) are precisely the stationary points of the formal sum

$$W(x) := \sum_{j \in \mathbb{Z}^d} S_j(x). \quad (1.6)$$

This follows because differentiation of (1.6) with respect to  $x_i$  produces exactly Eq. (1.4) and it explains why solutions to (1.4) are sometimes called *stationary* configurations.

In the case that the periodic onsite potential  $V(\xi)$  vanishes, the Frenkel–Kontorova equation (1.1) reduces to the discrete Laplace equation  $\Delta x = 0$ , for which it is easy to point out solutions. For instance, when  $\xi \in \mathbb{R}$  is an arbitrary number and  $\omega \in \mathbb{R}^d$  is an arbitrary vector, then the linear functions  $x^{\omega, \xi} : \mathbb{Z}^d \rightarrow \mathbb{R}$  defined by

$$x_i^{\omega, \xi} := \xi + \langle \omega, i \rangle$$

obviously satisfy  $\Delta x = 0$ . It moreover turns out that the  $x^{\omega, \xi}$  are *action-minimizers*, in the sense that for every finite subset  $B \subset \mathbb{Z}^d$  and every  $y : \mathbb{Z}^d \rightarrow \mathbb{R}$  with support in  $B$ , it holds that

$$\sum_{j \in \mathbb{Z}^d} (S_j(x^{\omega, \xi} + y) - S_j(x^{\omega, \xi})) \geq 0.$$

Note that this sum is actually finite and can be interpreted as  $W(x^{\omega, \xi} + y) - W(x^{\omega, \xi})$ .

**Definition 1.1.** Let  $x : \mathbb{Z}^d \rightarrow \mathbb{R}$  be a  $d$ -dimensional configuration. We say that  $\omega \in \mathbb{R}^d$  is the *rotation vector* of  $x$  if for all  $i \in \mathbb{Z}^d$ , the limit

$$\lim_{n \rightarrow \infty} \frac{x_{ni}}{n} \text{ exists and is equal to } \langle \omega, i \rangle.$$

Clearly, the rotation vector of  $x^{\omega, \xi}$  is equal to  $\omega$ . On the other hand, in dimension  $d \neq 1$ , a solution to (1.1) does not necessarily have a rotation vector. An example is the hyperbolic configuration  $x^h$  defined by  $x_i^h = i_1 i_2 \cdots i_{d-1} i_d$  which solves  $\Delta x = 0$ .

In Aubry–Mather theory, one is interested, among others, in answering the following questions: given a collection of local potentials  $S_j(x)$  satisfying the assumptions of Section 2, a number  $\xi \in \mathbb{R}$  and a vector  $\omega \in \mathbb{R}^d$ , does there always exist a solution  $x$  to Eq. (1.4) with rotation vector  $\omega$  and initial condition  $x_0 = \xi$ ? And if so, what is the structure of the solution set?

A rather complete answer to these questions is known. It turns out that solutions to (1.4) of all rotation vectors  $\omega \in \mathbb{R}^d$  exist. For example, it was shown by Bangert [2], that when  $\omega \in \mathbb{R}^d \setminus \mathbb{Q}^d$  is irrational, then there exists a unique nonempty collection of “recurrent” action-minimizers of rotation vector  $\omega$ . This is the Aubry–Mather set of rotation vector  $\omega$ . It is totally ordered, but may contain “gaps”. That is, given an arbitrary  $\xi \in \mathbb{R}$ , it may happen that the Aubry–Mather set of rotation vector  $\omega$  does not contain any configuration  $x$  satisfying the initial condition  $x_0 = \xi$ . It is known that in this case, the Aubry–Mather set is actually a Cantor set.

The basics of this classical theory will be reviewed in Sections 3 and 4 of this paper. In Section 3, we will study Birkhoff configurations, examples of which are the action-minimizing configurations of the Aubry–Mather sets. In Section 3.2, we will moreover prove some new results for  $d$ -dimensional periodic Birkhoff configurations. In Section 4, we will examine minimizing configurations and for completeness, we will reprove the classical result that global minimizers of every rotation vector exist and we will examine the properties of the Aubry–Mather set.

To investigate the existence of stationary configurations in the gaps of the Aubry–Mather sets, we propose to study the gradient flow of the formal action function, i.e. the flow of the differential equation

$$\frac{dx}{dt} = -\nabla W(x).$$

It was shown by Golé [7] that this flow is well defined on a suitable subspace  $\mathbb{X} \subset \mathbb{R}^{\mathbb{Z}^d}$  of configurations that contains all Birkhoff configurations. We will prove some regularity results for the gradient flow in Section 5 and we will discuss some of its qualitative properties in Section 6. The most notable of these is a strong monotonicity property or strong parabolic comparison principle, see Theorem 6.2.

The principal goal of this paper is then to prove the existence of a continuous one-dimensional gradient-flow invariant family of configurations that contains the Aubry–Mather set of rotation vector  $\omega$ . Such an interpolating family will be called a *ghost circle* and denoted  $\Gamma_\omega \subset \mathbb{R}^{\mathbb{Z}^d}$ . The precise definition of a ghost circle is given in Section 7.

Ghost circles were already constructed for twist maps by Golé [8]. Hence, they are well known to exist in dimension  $d = 1$ . Golé starts his construction by assuming that  $\omega = \frac{q}{p} \in \mathbb{Q}$  is rational and that an appropriate periodic action function  $W_{p,q}(x)$  is a Morse function. Under these assumptions, the existence of a periodic ghost circle follows from a combination of topological arguments and the parabolic comparison principle of the gradient flow. In Section 8.2, we will imitate the construction of these periodic Morse ghost circles in dimension  $d \neq 1$ . Each of these periodic ghost circles contains at least one global minimizer.

Our first main result is contained in Section 8.1. It generalizes results of Golé [9] on twist maps and it roughly states that for every rational  $\omega \in \mathbb{Q}^d$  and for every collection of potentials  $S_j(x)$ , one can find arbitrarily small perturbations of the  $S_j(x)$  that turn the periodic action  $W_{p,q}(x)$  into a Morse function. This statement is nontrivial in dimension  $d \neq 1$  and it holds because of group theoretic reasons that will be explained in Section 3.2.

The most important technical result of this paper is nevertheless a compactness theorem for ghost circles. It is presented in Section 9. It says that when the rotation vectors  $\omega_n$  converge to a rotation vector  $\omega_\infty$  and the local potentials  $S_j^n$  converge to potentials  $S_j^\infty$  and there exist ghost circles  $\Gamma_n$  for the potentials  $S_j^n$  of rotation vector  $\omega_n$ , then there is a ghost circle  $\Gamma_\infty$  for the potentials  $S_j^\infty$  of rotation vector  $\omega_\infty$ . Moreover, a subsequence of the  $\Gamma_n$  actually converges to  $\Gamma_\infty$  in a sense to be made precise. Together, all of the above shows that there are ghost circles of every rotation vector and for arbitrary potentials. Again, they contain at least one minimizer and hence the entire Aubry–Mather set of rotation vector  $\omega$ .

A similar compactness result was proved by Golé, see [9], for twist maps. The proof of this “monotone convergence theorem for ghost circles” relies on the fact that over time, two different solutions of the gradient flow must decrease their number of intersections. Hence, this proof is purely one-dimensional. Our proof, on the other hand, only depends on a quantitative version of the parabolic comparison principle, a so-called Harnack inequality. This inequality is stated and proved in Theorem 6.4.

As a consequence, we show in Section 10 that when the Aubry–Mather set is a Cantor set, then its gaps must either be completely foliated by minimizers, or contain at least one non-minimizing solution to (1.4).

## 2. Problem setup

Let us at this point introduce the generalized Frenkel–Kontorova lattice recurrence relations that we want to consider in this paper.

As was discussed before, we will assume that for all  $j \in \mathbb{Z}^d$  there is a function  $S_j$  that assigns a real value to every  $d$ -dimensional configuration:

$$S_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}.$$

These functions are required to have the conditions A–E described below and are called *local potentials*.

To formulate the first condition, let us assume that a finite subset  $B \subset \mathbb{Z}^d$  and an  $m \geq 0$  times continuously differentiable function  $s : \mathbb{R}^B \rightarrow \mathbb{R}$  are given. Then we can define a function  $S : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  by setting  $S(x) := s(x|_B)$ . This is just a way of saying that  $S$  depends only on the finitely many variables  $x_i$  for which  $i \in B$ .

Such an  $S$  has some convenient properties, most notably that it is continuous in the topology of pointwise convergence: if  $x^n, x^\infty \in \mathbb{R}^{\mathbb{Z}^d}$  and  $\lim_{n \rightarrow \infty} x^n = x^\infty$  pointwise, then obviously also  $\lim_{n \rightarrow \infty} S(x^n) = S(x^\infty)$ .

Moreover, it makes sense to speak of the partial derivatives of the function  $S$ : if  $j_1, \dots, j_k \in \mathbb{Z}^d$ , with  $0 \leq k \leq m$ , is a collection of lattice points, then the partial derivative  $\partial_{j_1, \dots, j_k} S : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  can simply be defined as

$$\partial_{j_1, \dots, j_k} S(x) := \begin{cases} (\partial_{j_1, \dots, j_k} S)(x|_B) & \text{if } j_1, \dots, j_k \in B, \\ 0 & \text{otherwise.} \end{cases}$$

These partial derivatives are also continuous with respect to pointwise convergence.

Finally, we recall the definition  $\|i\| := \sum_{k=1}^d |i_k|$ , for  $i \in \mathbb{Z}^d$ , and define  $B_j^r := \{k \in \mathbb{Z}^d \mid \|k - j\| \leq r\}$ . With all this in mind, we can formulate our first condition.

- A.** The functions  $S_j$  are twice continuously differentiable and of finite range. That is, there is an  $0 < r < \infty$  and for every  $j \in \mathbb{Z}^d$  there is a twice continuously differentiable function  $s_j : \mathbb{R}^{B_j^r} \rightarrow \mathbb{R}$  such that  $S_j(x) = s_j(x|_{B_j^r})$ .

In other words, the function  $S_j$  depends only on the finitely many variables  $x_k$  with  $\|k - j\| \leq r$ . Hence,  $S_j(x)$  has the interpretation of the “local energy” of the configuration  $x$  at lattice site  $j$  and we think of  $r$  as the finite range of the interaction.

To formulate condition B, it is convenient to introduce an action of  $\mathbb{Z}^d \times \mathbb{Z}$  on  $\mathbb{R}^{\mathbb{Z}^d}$  by “shifts”:

**Definition 2.1.** Let  $k \in \mathbb{Z}^d$  and  $l \in \mathbb{Z}$ . Then the shift operator  $\tau_{k,l} : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$  is defined by

$$(\tau_{k,l} x)_i := x_{i+k} + l.$$

Clearly, the graph of  $\tau_{k,l} x$ , viewed as a subset of  $\mathbb{Z}^d \times \mathbb{R}$ , is obtained by shifting the graph of  $x$  over the integer vector  $(-k, l)$ . This explains why the  $\tau_{k,l}$  are called *shift operators*.

- B.** The functions  $S_j$  are shift-invariant:  $S_j(\tau_{k,l} x) = S_{j+k}(x)$  for all  $j, k$  and  $l$ .

In fact, invariance of the  $S_j$  under  $\tau_{0,1}$  just means that  $S_j(x) = S_j(x + 1_{\mathbb{Z}^d})$  for all  $j$ , which means that  $S_j$  descends to a function on  $\mathbb{R}^{\mathbb{Z}^d} / \mathbb{Z}$ . Invariance of the  $S_j$  under the shifts  $\tau_{k,0}$  expresses the maximal spatial homogeneity of the local potentials. In fact, once one of the  $S_j$  is given, for instance  $S_0$ , then all the others are determined.

The next condition ensures the growth of the  $S_j$  at infinity:

- C.** The functions  $S_j$  are bounded from below and coercive in the following sense: for all  $k$  with  $\|k - j\| = 1$ ,

$$\lim_{|x_k - x_j| \rightarrow \infty} S_j(x) = \infty.$$

Condition C says that every function  $x \mapsto S_j(x)$  is as coercive as it can possibly be under the restriction that it satisfies the periodicity condition  $S_j(\tau_{0,1} x) = S_j(x)$ .

The following condition D is the most essential one:

- D.** The functions  $S_j$  satisfy the so-called *monotonicity condition*:

$$\partial_{i,k} S_j \leq 0 \quad \text{for all } j \text{ and all } i \neq k, \quad \text{while } \partial_{i,k} S_i < 0 \quad \text{for all } \|i - k\| = 1.$$

Condition D is also called a *twist condition* or *ferromagnetic condition*. It says that all mixed derivatives of the local potentials are non-positive, while some of them are strictly negative.

For technical reasons we will also assume:

**E.** The  $S_j$  have uniformly bounded second derivatives: there is a constant  $C$  such that

$$|\partial_{i,k} S_j| \leq C \quad \text{for all } i, j, k.$$

As in Section 1, we can now look for stationary configurations corresponding to these potentials.

**Definition 2.2.** A configuration  $x: \mathbb{Z}^d \rightarrow \mathbb{R}$  is called a *stationary point* for the local potentials  $S_j$  if for every finite subset  $B \subset \mathbb{Z}^d$  and every configuration  $y$  with support in its  $r$ -interior  $\hat{B}^{(r)} := \{i \in B \mid B_i^r \subset B\}$ , it holds that

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} W_B(x + \varepsilon y) = 0,$$

where  $W_B: \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  is defined as

$$W_B(x) := \sum_{j \in B} S_j(x). \quad (2.7)$$

In fact, by differentiating  $W_B$  with respect to an  $x_i$  with  $i \in \hat{B}^{(r)}$ , one obtains that  $x$  is a stationary point for the  $S_j$  if and only if it satisfies the variational monotone recurrence relation

$$\sum_{\|j-i\| \leq r} \partial_i S_j(x) = 0 \quad \text{for all } i \in \mathbb{Z}^d. \quad (2.8)$$

The goal of this paper is to find solutions of (2.8) and while doing so, we will exploit the variational principle that underlies it.

By the way, (2.8) is called *monotone* because condition C guarantees that the derivative of the left-hand side of (2.8) with respect to any of the  $x_k$  with  $k \neq i$ , is non-positive, while it is strictly negative if  $\|k-i\|=1$ .

Definition 2.2 moreover inspires the definition of a special type of solutions to (2.8):

**Definition 2.3.** A configuration  $x: \mathbb{Z}^d \rightarrow \mathbb{R}$  is called a *global minimizer* or *ground state* for the potentials  $S_j$  if for every finite subset  $B \subset \mathbb{Z}^d$  and every  $y: \mathbb{Z}^d \rightarrow \mathbb{R}$  with support in  $\hat{B}^{(r)}$ ,

$$W_B(x+y) - W_B(x) = \sum_{j \in \mathbb{Z}^d} (S_j(x+y) - S_j(x)) \geq 0.$$

Clearly, global minimizers are automatically stationary and hence satisfy the recurrence relation (2.8).

**Example 2.4.** It is easy to check that the Frenkel–Kontorova potentials given in (1.5) satisfy conditions A–E. In fact, the range of interaction is  $r=1$ , and  $\partial_{j,k} S_j = -\frac{1}{4d}$  for  $\|j-k\|=1$ .

In the particular case that  $V(\xi) \equiv 0$  all solutions of (1.1) are actually global minimizers. This follows because every  $y \mapsto W_B(x+y)$  is strictly convex if  $V(\xi) \equiv 0$  and hence only has one stationary point, which is minimizing.

### 3. Spaces of configurations

In this section, we introduce certain spaces of configurations that are often encountered in classical Aubry–Mather theory. We will moreover study some of their properties. Most of the definitions and results in this section are standard, but to the best of our knowledge Lemma 3.10 and Theorem 3.12 in Section 3.2 are new. We start by recalling the following definition:

**Definition 3.1.** Let  $x : \mathbb{Z}^d \rightarrow \mathbb{R}$  be a  $d$ -dimensional configuration. We say that  $\omega = \omega(x) \in \mathbb{R}^d$  is the *rotation vector* of  $x$  if for all  $i \in \mathbb{Z}^d$ , the limit

$$\lim_{n \rightarrow \infty} \frac{x_{ni}}{n} \text{ exists and is equal to } \langle \omega, i \rangle.$$

The space of configurations with rotation vector  $\omega$  is denoted

$$\mathbb{X}_\omega := \{x : \mathbb{Z}^d \rightarrow \mathbb{R} \mid \omega(x) = \omega\}.$$

#### 3.1. Birkhoff configurations

We will now introduce the concept of a well-ordered lattice configuration.

**Definition 3.2.** On the configuration space  $\mathbb{R}^{\mathbb{Z}^d}$  we define the relations  $\leq$ ,  $<$  and  $\ll$  by:

- $x \leq y$  if  $x_i \leq y_i$  for every  $i \in \mathbb{Z}^d$ .
- $x < y$  if  $x \leq y$ , but  $x \neq y$ .
- $x \ll y$  if  $x_i < y_i$  for every  $i \in \mathbb{Z}^d$ .

Similarly for  $\geq$ ,  $>$  and  $\gg$ .

Recall the definition of the shift operators  $\tau_{k,l} : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$ . The partial orderings defined above, now allow us to make the following definition, as in for instance [3] and [11].

**Definition 3.3.** A configuration  $x \in \mathbb{R}^{\mathbb{Z}^d}$  is called a *Birkhoff configuration* or a *well-ordered configuration*, if for all  $k \in \mathbb{Z}^d$  and  $l \in \mathbb{Z}$ ,

$$\text{either } \tau_{k,l}x \geq x \text{ or } \tau_{k,l}x \leq x. \quad (3.9)$$

Definition 3.3 says that the graph of a Birkhoff configuration  $x$  does not cross any of its integer translates. The space of Birkhoff configurations will be denoted  $\mathcal{B} \subset \mathbb{R}^{\mathbb{Z}^d}$  and it inherits the topology of pointwise convergence. Birkhoff configurations will play an essential role in the remainder of this paper. Birkhoff configurations of every rotation vector exist: for every  $\omega \in \mathbb{R}^d$  the linear configuration  $x^\omega$  defined by  $x_i^\omega := \langle \omega, i \rangle$  is an example.

**Remark 3.4.** When  $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  is an orientation preserving circle homeomorphism, then it admits a *lift* to a strictly increasing map  $H : \mathbb{R} \rightarrow \mathbb{R}$  that satisfies  $H(\xi + 1) = H(\xi) + 1$  and  $H(\xi) \bmod 1 = h(\xi \bmod 1)$ .

Let us now denote by  $x(\xi) : \mathbb{Z} \rightarrow \mathbb{R}$  the  $H$ -orbit of  $\xi \in \mathbb{R}$ , defined by  $x(\xi)_i := H^i(\xi)$ . Then it is clear that for  $\xi_1, \xi_2 \in \mathbb{R}$  with  $\xi_1 < \xi_2$ , one has that  $x(\xi_1) \ll x(\xi_2)$ . In turn this implies that each  $x(\xi)$  is a Birkhoff sequence. Thus, ordering is a very natural concept in the theory of circle homeomorphisms.

The following result is folklore and it goes back to Poincaré, who proved it in the case  $d = 1$  and in the context of circle homeomorphisms, for which it implies that circle homeomorphisms have a unique rotation number.

For  $d = 1$ , the proof of Lemma 3.5 can be found for instance in [9]. For completeness, we include the proof for  $d > 1$  here. Lemma 3.5 says that the graph of a Birkhoff configuration  $x$  lies uniformly close to the graph of the affine configuration  $i \mapsto x_0 + \langle \omega, i \rangle$ .

**Lemma 3.5.** *Let  $x \in \mathbb{R}^{\mathbb{Z}^d}$  be a Birkhoff configuration. Then  $x$  has a rotation vector  $\omega = \omega(x)$  and*

$$|x_i - x_0 - \langle \omega(x), i \rangle| \leq 1.$$

Moreover, the map  $x \mapsto \omega(x)$ ,  $\mathcal{B} \rightarrow \mathbb{R}^d$  is continuous with respect to pointwise convergence. We write

$$\mathcal{B}_\omega := \{x \in \mathcal{B} \mid \omega(x) = \omega\}.$$

**Proof.** We will assume that the result is true for  $d = 1$  and we choose  $i, j \in \mathbb{Z}^d$ . Then the sequence  $n \mapsto x_{ni+j}$  is a one-dimensional Birkhoff sequence and hence its rotation number  $\omega_{i,j}$  exists and is equal to  $\lim_{n \rightarrow \infty} \frac{x_{ni+j} - x_0}{n}$ . Moreover,  $|x_{ni+j} - x_j - \langle \omega_{i,j}, n \rangle| \leq 1$ . We first of all remark that  $\omega_{i,j}$  does not depend on  $j$ , and hence can be denoted  $\omega_i$ . This follows because the Birkhoff property of  $x$  ensures that the sequences  $n \mapsto x_{ni+j} = (\tau_{j,0}x)_{ni}$  and  $n \mapsto x_{ni}$  do not cross. Now denote by  $e_1 = (1, 0, \dots, 0)$ ,  $e_2 = (0, 1, 0, \dots, 0)$ , etc. the standard basis of  $\mathbb{Z}^d$  and define  $\omega := (\omega_{e_1}, \dots, \omega_{e_d})$ . Then,

$$\begin{aligned} |x_i - x_0 - \langle \omega, i \rangle| &= |x_{(i_1, i_2, \dots, i_d)} - x_{(0, i_2, \dots, i_d)} - i_1 \omega_{e_1} + \dots + x_{(0, \dots, 0, i_d)} - x_0 - i_d \omega_{e_d}| \\ &\leq |x_{(i_1, i_2, \dots, i_d)} - x_{(0, i_2, \dots, i_d)} - i_1 \omega_{e_1}| + \dots + |x_{(0, \dots, 0, i_d)} - x_0 - i_d \omega_{e_d}| \leq d. \end{aligned}$$

This clearly implies that  $\lim_{n \rightarrow \infty} \frac{x_{ni}}{n} = \langle \omega, i \rangle$ , while the Birkhoff property of the sequence  $n \mapsto x_{ni}$  then implies that in fact,  $|x_i - x_0 - \langle \omega, i \rangle| \leq 1$ .

The continuity of  $x \mapsto \omega(x)$  follows immediately from the continuity in the one-dimensional case.  $\square$

The following proposition is equally standard. In particular, it will allow us to take limits of Birkhoff configurations with rational rotation vectors in order to produce Birkhoff configurations with irrational rotation vectors.

Recall the action  $\tau_{0,1} : x \mapsto x + 1$  on  $\mathbb{R}^{\mathbb{Z}^d}$ . It can be used to identify sequences that differ by an integer. The quotient space is denoted  $\mathbb{R}^{\mathbb{Z}^d} / \mathbb{Z}$ . Note that every element  $[x]$  in this quotient space has a unique representative  $x$  with  $x_0 \in [0, 1)$ .

**Proposition 3.6.** *Let  $K \subset \mathbb{R}^d$  be compact and let  $\mathcal{B}_K := \bigcup_{\omega \in K} \mathcal{B}_\omega$ . Then  $\mathcal{B}_K / \mathbb{Z}$  is compact in the topology of pointwise convergence.*

**Proof.** By definition,  $\mathcal{B}$  is closed in the topology of pointwise convergence. Moreover, by Proposition 3.5,  $\mathcal{B}_K / \mathbb{Z}$  is a closed subset of

$$\{[x] \in \mathbb{R}^{\mathbb{Z}^d} / \mathbb{Z} \mid x_k = x_0 + k \cdot \omega + y_k \text{ with } ([x]_0, \omega, y_k) \in \mathbb{R} / \mathbb{Z} \times K \times [-1, 1]^{\mathbb{Z}^d}\},$$

which is compact in the topology of pointwise convergence. This follows from Tychonov's theorem.  $\square$

The following corollary of the compactness of  $\mathcal{B}_K / \mathbb{Z}$  is trivial, but it has important implications.



**Corollary 3.7.** Let  $K \subset \mathbb{R}^d$  be compact and let  $B \subset \mathbb{Z}^d$  be a finite subset. Assume that  $s : \mathbb{R}^B / \mathbb{Z} \rightarrow \mathbb{R}$  is a continuous function. Then  $S : \mathbb{R}^{\mathbb{Z}^d} / \mathbb{Z} \rightarrow \mathbb{R}$  defined by  $S(x) = s(x|_B)$  attains its maximum and minimum values on  $\mathcal{B}_K / \mathbb{Z}$ .

**Proof.** This follows because such an  $S$  is continuous with respect to pointwise convergence and  $\mathcal{B}_K / \mathbb{Z}$  is compact.  $\square$

Applied to  $S = \partial_{i,k} S_i$  with  $\|i - k\| = 1$ , and recalling the twist condition  $\partial_{i,k} S_i < 0$ , Corollary 3.7 implies that there is a  $\lambda > 0$  such that  $\partial_{i,k} S_i(x) < -\lambda < 0$  for all  $x \in \mathcal{B}_K$ . In other words, the twist condition D is automatically uniform on  $\mathcal{B}_K$ . Similarly, even if one does not impose condition E, there is a constant  $C > 0$  such that  $|\partial_{i,k} S_j| \leq C$  for all  $i, k$  and  $j$ , uniformly on  $\mathcal{B}_K$ .

We finish this section with a simple and well-known proposition that expresses that the number theoretical properties of the rotation vector  $\omega$  of a Birkhoff configuration  $x$  decide to a large extent whether  $\tau_{k,l}x > x$  or  $\tau_{k,l}x < x$ .

**Proposition 3.8.** Let  $\omega \in \mathbb{R}^d$  and  $x \in \mathcal{B}_\omega$ . If  $\langle \omega, k \rangle + l > 0$ , then  $\tau_{k,l}x > x$  and if  $\langle \omega, k \rangle + l < 0$ , then  $\tau_{k,l}x < x$ .

**Proof.** Denote by  $x^\omega \in \mathcal{B}_\omega$  the linear configuration defined by  $x_i^\omega := \langle \omega, i \rangle$ . Then  $\tau_{k,l}x^\omega - x^\omega = \langle k, \omega \rangle + l$ . Suppose for instance that  $\langle k, \omega \rangle + l > 0$ , that is that  $\tau_{k,l}x^\omega \gg x^\omega$ , but assume on the other hand that  $\tau_{k,l}x \leq x$ . This means that  $\tau_{k,l}x - x \leq 0$  and hence also that  $\tau_{k,l}^2x - \tau_{k,l}x = \tau_{k,0}(\tau_{k,l}x - x) \leq 0$ . Thus,  $\tau_{k,l}^2x - x = (\tau_{k,l}^2x - \tau_{k,l}x) + (\tau_{k,l}x - x) \leq 0$ , i.e.  $\tau_{k,l}^2x \leq x$ . By induction we then find that  $\tau_{k,l}^n x \leq x$ , for every  $n \geq 1$ . On the other hand,  $\tau_{k,l}^n x^\omega = x^\omega + n(\langle k, \omega \rangle + l)$ . This contradicts the fact that  $\sup_i |\tau_{k,l}^n(x^\omega - x)_i| = \sup_i |(x^\omega - x)_i| \leq |x_0| + 2$  is uniformly bounded in  $n$ .  $\square$

### 3.2. Periodicity

It turns out convenient to consider periodic configurations. To define these, let  $p_1, \dots, p_c \in \mathbb{Z}^d$  be  $0 \leq c \leq d$  linearly independent integer vectors and let  $q_1, \dots, q_c \in \mathbb{Z}$  be  $c$  integers. Then we set

$$\mathbb{X}_{p,q} := \{x : \mathbb{Z}^d \rightarrow \mathbb{R} \mid \tau_{p_j,q_j}x = x \text{ for all } j = 1, \dots, c\}.$$

We say that a configuration  $x \in \mathbb{X}_{p,q}$  is *periodic* with periods  $(p_1, q_1), \dots, (p_c, q_c)$ . The collection of periods of  $\mathbb{X}_{p,q}$  is a lattice of rank  $c$ , that we denote by

$$J_{p,q} := \left\{ \sum_{j=1}^c m_j(p_j, q_j) \mid m_j \in \mathbb{Z} \right\} \subset \mathbb{Z}^d \times \mathbb{Z}.$$

An element of  $\mathbb{X}_{p,q}$  can have a rotation vector, but this rotation vector cannot be arbitrary: when  $x : \mathbb{Z}^d \rightarrow \mathbb{R}$  is a configuration of rotation vector  $\omega$  and  $\tau_{p_j,q_j}x = x$ , then  $x_{np_j} = x_0 - nq_j$ , so that  $\lim_{n \rightarrow \infty} \frac{x_{np_j}}{n} = -q_j$ , that is

$$\langle \omega, p_j \rangle + q_j = 0 \text{ when } \mathbb{X}_{p,q} \cap \mathbb{X}_\omega \neq \emptyset.$$

Another way to express this is that when  $\mathbb{X}_{p,q} \cap \mathbb{X}_\omega \neq \emptyset$ , then  $J_{p,q} \subset I_\omega$ , where the lattice  $I_\omega$  is defined as

$$I_\omega := \{(k, l) \in \mathbb{Z}^d \times \mathbb{Z} \mid \langle k, \omega \rangle + l = 0\} \subset \mathbb{Z}^d \times \mathbb{Z}.$$

On the other hand, when  $x$  has rotation vector  $\omega$  and  $\langle \omega, k \rangle + l = 0$ , then this does not imply that  $\tau_{k,l}x = x$ . We therefore define

$$\bar{\mathbb{X}}_\omega := \{x \in \mathbb{X}_\omega \mid \tau_{k,l}x = x \text{ when } \langle \omega, k \rangle + l = 0\}.$$

The elements of  $\bar{\mathbb{X}}_\omega$  are called *maximally periodic* as they have all the periods that an element of  $\mathbb{X}_\omega$  can possibly have.  $\bar{\mathbb{X}}_\omega$  is nonempty because it contains the linear configuration  $x^\omega$  defined by  $x_i^\omega = \langle \omega, i \rangle$ . This is true because  $\tau_{k,l}x^\omega = x^\omega + \langle \omega, k \rangle + l$ .

**Definition 3.9.** A  $\mathbb{Z}$ -basis  $(p_1, q_1), \dots, (p_c, q_c)$  of  $I_\omega$  is called a collection of *principal periods* for  $\omega$ . That is,  $(p_1, q_1), \dots, (p_c, q_c)$  are principal periods for  $\omega$  if and only if  $\mathbb{X}_{p,q} \cap \mathbb{X}_\omega = \bar{\mathbb{X}}_\omega$ .

Of course, a set of principal periods for  $\omega \in \mathbb{R}^d$  always exists, but it is not unique.

At this point, let us make some group theoretic remarks. First of all, we remind the reader that we can think of the shift operators  $\tau_{k,l}$  as defining a group action of  $\mathbb{Z}^d \times \mathbb{Z}$  on the space of configurations:

$$\tau : (\mathbb{Z}^d \times \mathbb{Z}) \times \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}^{\mathbb{Z}^d}, \quad ((k, l), x) \mapsto \tau_{k,l}x.$$

Clearly, because  $\mathbb{Z}^d \times \mathbb{Z}$  is Abelian, when  $\tau_{p_j, q_j}x = x$ , then also  $\tau_{p_j, q_j}(\tau_{k,l}x) = \tau_{k,l}x$ , and thus  $\tau$  leaves  $\mathbb{X}_{p,q}$  invariant. Moreover, because the elements of  $J_{p,q}$  fix all elements of  $\mathbb{X}_{p,q}$ , we have that when  $x \in \mathbb{X}_{p,q}$  and  $(k, l) = (K, L) + \sum_j m_j(p_j, q_j)$  for certain integers  $m_j$ , then  $\tau_{k,l}x = \tau_{K,L}x$ . This shows that  $\tau$  induces an action of  $(\mathbb{Z}^d \times \mathbb{Z})/J_{p,q}$  on  $\mathbb{X}_{p,q}$ . We recall that this action is called *free* if for every  $(k, l) \notin J_{p,q}$  and every  $x \in \mathbb{X}_{p,q}$  it holds that  $\tau_{k,l}x \neq x$ . We now have the following quite obvious characterization of  $\bar{\mathbb{X}}_\omega$ :

**Lemma 3.10.** Assume that  $\mathbb{X}_{p,q} \cap \bar{\mathbb{X}}_\omega \neq \emptyset$ . Then the  $\tau$ -action of  $(\mathbb{Z}^d \times \mathbb{Z})/J_{p,q}$  on  $\mathbb{X}_{p,q} \cap \bar{\mathbb{X}}_\omega$  is free if and only if the  $(p_j, q_j)$  are principal periods for  $\omega$ , i.e. if and only if  $\mathbb{X}_{p,q} \cap \bar{\mathbb{X}}_\omega = \bar{\mathbb{X}}_\omega$ .

**Proof.** Let us start by assuming that the  $(p_j, q_j)$  are principal periods for  $\omega$ , that is that  $J_{p,q} = I_\omega$ . We want to show that then the action of  $(\mathbb{Z}^d \times \mathbb{Z})/I_\omega$  on  $\mathbb{X}_{p,q} \cap \bar{\mathbb{X}}_\omega$  is free. But a nontrivial equivalence class in  $(\mathbb{Z}^d \times \mathbb{Z})/I_\omega$  is represented by an element  $(k, l)$  with  $\langle \omega, k \rangle + l \neq 0$  and it is clear that this inequality implies that  $\tau_{k,l}x \neq x$  if  $x$  has rotation vector  $\omega$ .

In the other direction, suppose the action is not free. Then there is a  $(k, l) \notin J_{p,q}$  and an  $x \in \mathbb{X}_{p,q} \cap \bar{\mathbb{X}}_\omega$  with  $\tau_{k,l}x = x$ . Clearly, such  $(k, l)$  must satisfy  $\langle \omega, k \rangle + l = 0$ , that is  $(k, l) \in I_\omega$ . Thus,  $J_{p,q} \neq I_\omega$ .  $\square$

The case that  $\omega \in \mathbb{Q}^d$  is especially nice. We have the following:

**Proposition 3.11.**  $\omega \in \mathbb{Q}^d$  if and only if  $I_\omega$  has rank  $d$ . When  $(p_1, q_1), \dots, (p_d, q_d) \in I_\omega$  are linearly independent, then  $\mathbb{X}_{p,q}$  is finite-dimensional and  $\mathbb{X}_{p,q} \subset \bar{\mathbb{X}}_\omega$ . In particular, when  $(p_1, q_1), \dots, (p_d, q_d)$  are principal periods, then  $\mathbb{X}_{p,q} = \bar{\mathbb{X}}_\omega$ .

**Proof.** Let us suppose that  $\omega \in \mathbb{Q}^d$ , for instance  $\omega = (\frac{a_1}{b_1}, \dots, \frac{a_d}{b_d})$  for integers  $a_j$  and  $b_j$ . Then  $\langle \omega, P_j \rangle + Q_j = 0$  for  $P_j := (0, \dots, 0, b_j, 0, \dots, 0)$  and  $Q_j = -a_j$ . This shows that  $I_\omega$  has rank  $d$ .

On the other hand, when  $I_\omega$  has rank  $d$ , then we can choose linearly independent  $(p_1, q_1), \dots, (p_d, q_d) \in I_\omega$ . If we now denote by  $p$  the  $d \times d$ -matrix with integer coefficients  $(p_1, \dots, p_d)$  and by  $q = (q_1, \dots, q_d) \in \mathbb{Z}^d$  the integer vector of length  $d$ , then we can write the equations  $\langle p_j, \omega \rangle + q_j = 0$  as the matrix equality  $p^T \omega + q = 0$ , where  $p^T$  denotes the transpose of the matrix  $p$ . This implies that the final column in the rank- $d$  matrix  $(p^T, q)$  is degenerate, i.e. that  $p^T$  is invertible. In particular,  $\omega = -p^{-T}q \in \mathbb{Q}^d$ , with  $p^{-T}$  the inverse transpose of the matrix  $p$ .

Moreover, the fact that  $p$  is invertible implies that  $\mathbb{X}_{p,q}$  is finite-dimensional. More precisely, let us define

$$B_p := p([0, 1)^d) \cap \mathbb{Z}^d.$$

Then  $B_p$  is a fundamental domain for  $p$ , that is for every  $i \in \mathbb{Z}^d$  there is a unique  $k \in B_p$  with  $k = i \bmod p(\mathbb{Z}^d)$ . It is not hard to show that this implies that the map  $x \mapsto x|_{B_p}$  from  $\mathbb{X}_{p,q}$  to  $\mathbb{R}^{B_p}$  is an isomorphism.

Thus, we have that  $\dim \mathbb{R}^{B_p} = |B_p| = \text{vol}_d(p[0, 1)^d) = |\det p|$  and hence  $\mathbb{X}_{p,q}$  is finite-dimensional.

In turn this implies that  $\mathbb{X}_{p,q} \subset \mathbb{X}_\omega$ , because any  $x \in \mathbb{X}_{p,q}$  satisfies  $\sup_{i \in \mathbb{Z}^d} \{x_i - \langle \omega, i \rangle\} = \sup_{i \in B_p} \{x_i - \langle \omega, i \rangle\} < \infty$  and the configuration  $i \mapsto \langle \omega, i \rangle$  has rotation vector  $\omega$ .

If  $(p_1, q_1), \dots, (p_d, q_d)$  are principal periods, then the above implies that  $\bar{\mathbb{X}}_\omega = \mathbb{X}_\omega \cap \mathbb{X}_{p,q} = \mathbb{X}_{p,q}$ .  $\square$

After these general considerations, let us now return to the Birkhoff configurations defined in Section 3.1. Let us denote the set of maximally periodic Birkhoff configurations of rotation vector  $\omega$  by

$$\bar{\mathcal{B}}_\omega := \{x \in \mathcal{B}_\omega \mid \tau_{k,l}x = x \text{ when } \langle k, \omega \rangle + l = 0\}.$$

The following theorem expresses that periodic Birkhoff configurations are automatically maximally periodic.

**Theorem 3.12.** *Let  $\omega \in \mathbb{R}^d$ , denote  $c := \text{rank}_{\mathbb{Z}}(I_\omega)$  and let  $(p_1, q_1), \dots, (p_c, q_c) \in I_\omega$  be linearly independent. Then*

$$\mathbb{X}_{p,q} \cap \mathcal{B}_\omega = \bar{\mathcal{B}}_\omega.$$

**Proof.** Let  $x \in \mathbb{X}_{p,q} \cap \mathcal{B}_\omega$ , that is  $x$  is Birkhoff, has rotation vector  $\omega$  and  $\tau_{p_j, q_j}x = x$  for all  $j = 1, \dots, c$ . We need to show that whenever  $\langle \omega, k \rangle + l = 0$ , then  $\tau_{k,l}x = x$ . So let us assume that  $\tau_{k,l}x \neq x$ . Because  $x$  is Birkhoff, we may assume that  $\tau_{k,l}x > x$ : the case  $\tau_{k,l}x < x$  is similar. This assumption implies that  $\tau_{nk, nl}x = \tau_{k,l}^n x > x$  as well, for every  $n \geq 1$ . We claim that this is not possible.

To prove this claim, we remark that there must be an  $n \in \mathbb{N}$  and  $m_1, \dots, m_c \in \mathbb{Z}$  so that  $n(k, l) = \sum_j m_j(p_j, q_j)$ . This is because by assumption the  $(p_j, q_j)$  span a sublattice of  $I_\omega$  of maximal rank. We therefore have that  $\tau_{nk, nl}x = (\tau_{p_1, q_1}^{m_1} \circ \dots \circ \tau_{p_c, q_c}^{m_c})x = x$ . This is a contradiction and hence,  $\tau_{k,l}x = x$ .  $\square$

In dimension  $d = 1$ , Theorem 3.12 simply says that a Birkhoff configuration of period  $(np, nq)$  automatically has period  $(p, q)$ . That is, the period  $(p, q)$  of a one-dimensional Birkhoff configuration can be chosen relatively prime. Theorem 3.12 is the  $d$ -dimensional variant of this statement.

In spite of Theorem 3.12, it should be remarked that in general,  $\bar{\mathcal{B}}_\omega \neq \mathcal{B}_\omega$ , that is not all Birkhoff configurations of rotation vector  $\omega$  are periodic. Counterexamples are easy to find.

#### 4. Classical Aubry–Mather theory

We are now ready to discuss the most well-known results of classical Aubry–Mather theory in the context of lattice equations. The concepts and results of this section are widely known, but we chose to present them in a perhaps slightly unconventional manner.

##### 4.1. Fully periodic minimizers

Throughout Section 4.1, we will assume that  $\omega \in \mathbb{Q}^d$  and  $(p_1, q_1), \dots, (p_d, q_d) \in \mathbb{Z}^d \times \mathbb{Z}$  are linearly independent.

We are interested in solutions to (2.8) that lie in  $\mathbb{X}_{p,q}$ . We start by noting the presence of a variational structure. Recalling the definition  $B_p = \mathbb{Z}^d \cap p([0, 1)^d)$ , one has

**Proposition 4.1.** *A configuration  $x \in \mathbb{X}_{p,q}$  solves (2.8) if and only if it is a stationary point of the periodic action function*

$$W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R} \quad \text{defined by} \quad W_{p,q}(x) := W_{B_p}(x) = \sum_{j \in B_p} S_j(x), \quad (4.10)$$

with respect to variations in  $\mathbb{X}_{p,q}$ .

**Proof.** We start by recalling the shift-invariance of the local potentials, condition B, which says that  $S_{j+k}(x) = S_j(\tau_{k,l}x)$  for all  $k$  and  $l$  and all  $x \in \mathbb{R}^{\mathbb{Z}^d}$ . Differentiation of this identity with respect to  $x_i$  then gives that  $\partial_i S_{j+k}(x) = \partial_{i-k} S_j(\tau_{k,l}x)$ . These equalities respectively imply that for  $x \in \mathbb{X}_{p,q}$  it holds that  $S_{j+pk}(x) = S_j(x)$  and  $\partial_i S_{j+pk}(x) = \partial_{i-pk} S_j(x)$  for all  $k \in \mathbb{Z}^d$ .

Now let  $x \in \mathbb{X}_{p,q}$ , choose an  $i \in \mathbb{Z}^d$  and define  $e_i \in \mathbb{X}_{p,0}$  by letting  $(e_i)_j = 1$  if  $j = i \bmod p(\mathbb{Z}^d)$  and  $(e_i)_j = 0$  otherwise. Then  $x + e_i \in \mathbb{X}_{p,q}$  and

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} W_{p,q}(x + \varepsilon e_i) = \sum_{k \in \mathbb{Z}^d} \sum_{j \in B_p} \partial_{i+pk} S_j(x) = \sum_{k \in \mathbb{Z}^d} \sum_{j \in B_p} \partial_i S_{j-pk}(x) = \sum_{j \in \mathbb{Z}^d} \partial_i S_j(x).$$

Of course all these sums are finite.  $\square$

Note that  $W_{p,q} = W_{B_p}$  is actually well defined for any  $x \in \mathbb{R}^{\mathbb{Z}^d}$ , but in this section, we restrict it to a function on  $\mathbb{X}_{p,q}$ . As such, it is shift-invariant:

**Lemma 4.2** (Shift-invariance). *The function  $W_{p,q}$  is  $\tau$ -invariant: for  $x \in \mathbb{X}_{p,q}$  and  $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}$  arbitrary, it holds that  $W_{p,q}(\tau_{k,l}x) = W_{p,q}(x)$ .*

**Proof.** In general,  $S_{j+k}(x) = S_j(\tau_{k,l}x)$ , so that if  $x \in \mathbb{X}_{p,q}$ , then  $S_{j+pk}(x) = S_j(x)$  for all  $k \in \mathbb{Z}^d$ . Thus,  $W_{p,q}(\tau_{k,l}x) = \sum_{j \in B_p} S_j(\tau_{k,l}x) = \sum_{j \in B_p} S_{j+k}(x) = \sum_{j \in k+B_p} S_j(x) = \sum_{j \in B_p} S_j(x) = W_{p,q}(x)$ . The fourth equality follows as both  $B_p$  and  $k + B_p$  are fundamental domains of  $\mathbb{Z}^d/p(\mathbb{Z}^d)$ , so that for every  $j \in k + B_p$  there is a unique  $i \in B_p$  for which  $i = j \bmod p(\mathbb{Z}^d)$ .  $\square$

**Theorem 4.3** (Existence). *The action  $W_{p,q}$  attains its minimum on  $\mathbb{X}_{p,q}$ .*

**Proof.** Since  $W_{p,q}(\tau_{0,1}x) = W_{p,q}(x)$  for  $x \in \mathbb{X}_{p,q}$ , clearly  $W_{p,q}$  descends to a function on  $\mathbb{X}_{p,q}/\mathbb{Z}$ . Every element in this quotient space has a representative  $x$  with  $x_0 \in [0, 1]$ .

Choose a cube  $C_N = \{i \in \mathbb{Z}^d \mid |i_k| \leq N \text{ for all } k = 1, \dots, d\}$  that contains  $B_p$  and choose a  $k \in \mathbb{Z}^d$  and an  $n \in \mathbb{N}$  such that  $k + nB_p$  in turn contains  $C_N$ . Moreover, remember that  $W_{k+nB_p} = n^d W_{p,q}$  on  $\mathbb{X}_{p,q}$ .

The coercivity of the  $S_i$ , condition C, implies that for all  $j$  with  $\|j\| = 1$ , it holds that for  $x \in \mathbb{X}_{p,q}$  with  $x_0 \in [0, 1]$ , we have that  $\lim_{|x_j| \rightarrow \infty} W_{k+nB_p}(x) = \infty$ . And hence by induction that for all  $j \in C_N$  and  $x \in \mathbb{X}_{p,q}$  with  $x_0 \in [0, 1]$ , it holds that  $\lim_{|x_j| \rightarrow \infty} W_{k+nB_p}(x) = \infty$ . Because  $B_p \subset C_N$  and  $W_{p,q} = n^{-d} W_{k+nB_p}$ , this means in particular that for all  $j \in B_p$  and  $x \in \mathbb{X}_{p,q}$  with  $x_0 \in [0, 1]$  it holds that  $\lim_{|x_j| \rightarrow \infty} W_{p,q}(x) = \infty$ . Hence,  $W_{p,q}$  attains its minimum on  $\mathbb{X}_{p,q}$ .  $\square$

The configurations that minimize  $W_{p,q}$  on  $\mathbb{X}_{p,q}$  will be called  $p, q$ -minimizers. Note that other extremal points of  $W_{p,q}$  in  $\mathbb{X}_{p,q}$ , such as saddle points, may also exist. Under certain mild conditions their existence will be proved later in this paper.

The following lemma is well known. We took the proof from [5].

**Lemma 4.4** (Minimum–maximum property). Assume the periodic configurations  $x, y \in \mathbb{X}_{p,q}$  are  $p, q$ -minimizers. Then also  $m := \min\{x, y\}$  and  $M := \max\{x, y\}$  are  $p, q$ -minimizers.

**Proof.** It is obvious that  $m, M \in \mathbb{X}_{p,q}$ . Write  $\alpha := M - x$  and  $\beta := m - x$  and observe that  $\alpha > 0, \beta < 0$ , while  $\text{supp}(\alpha) \cap \text{supp}(\beta) = \emptyset$  and  $y = M + m - x = \alpha + m = \alpha + \beta + x$ . The proof is done, if we show that

$$W_{p,q}(x) + W_{p,q}(y) \geq W_{p,q}(M) + W_{p,q}(m).$$

This is the same as showing

$$W_{p,q}(x + \alpha + \beta) - W_{p,q}(x + \alpha) - W_{p,q}(x + \beta) + W_{p,q}(x) \geq 0.$$

The left-hand side of this inequality can be put in integral form as

$$\int_0^1 \int_0^1 \frac{\partial^2}{\partial t \partial s} W_{p,q}(x + \alpha t + \beta s) ds dt = \sum_{i,k \in \mathbb{Z}^d} \sum_{j \in B_p} \left( \int_0^1 \int_0^1 \partial_{i,k} S_j(x + t\alpha + s\beta) ds dt \right) \alpha_i \beta_k.$$

Since  $\text{supp}(\alpha) \cap \text{supp}(\beta) = \emptyset$ , we have that  $\alpha_i \beta_i = 0$  for all  $i$ . Moreover, the twist condition  $\partial_{i,k} S_j \leq 0$  for all  $i \neq k$  and the inequalities  $\alpha_i \beta_k \leq 0$ , guarantee that the remaining terms in the sum are non-negative.  $\square$

This is now used to prove the following famous lemma:

**Lemma 4.5** (Aubry's lemma). Assume the configurations  $x \neq y \in \mathbb{X}_{p,q}$  are  $p, q$ -minimizers. Then either  $x \ll y$  or  $y \ll x$ .

**Proof.** We pursue a proof by contradiction. Denote again  $m := \min\{x, y\}$ . Suppose that for instance that  $m < x$  but that is not true that  $m \ll x$ . The case that  $m < y$  and not  $m \ll y$  is similar. The assumption implies that there are indices  $i, k \in \mathbb{Z}^d$  with  $\|i - k\| = 1$  such that  $m_i = x_i$  and  $m_k < x_k$ . Now we compute

$$\begin{aligned} \sum_{j \in \mathbb{Z}^d} (\partial_i S_j(x) - \partial_i S_j(m)) &= \sum_{j \in \mathbb{Z}^d} \int_0^1 \frac{d}{dt} \partial_i S_j(tx + (1-t)m) dt \\ &= \sum_{j, l \in \mathbb{Z}^d} \left( \int_0^1 \partial_{i,l} S_j(tx + (1-t)m) dt \right) (x_l - m_l). \end{aligned}$$

Recall that  $x_i = m_i$ , while, by the twist condition, for every  $l \neq i$ , it holds that  $\partial_{i,l} S_j \leq 0$  and  $x_l - m_l \geq 0$ . Thus, every term in the above sum is non-positive. But for the  $k$  chosen above,  $\partial_{i,k} S_i < 0$ , while  $x_k - m_k > 0$ . This proves that  $\sum_j \partial_i S_j(x) \neq \sum_j \partial_i S_j(m)$ . This contradicts the fact that, by the lemma above, both  $m$  and  $x$  are  $p, q$ -minimizers and must therefore both be stationary.  $\square$

**Corollary 4.6.** Periodic minimizers are Birkhoff configurations.

**Proof.** Let  $x \in \mathbb{X}_{p,q}$  be a minimizer. Then for any  $k \in \mathbb{Z}^d$  and  $l \in \mathbb{Z}$ , we have that  $W_{p,q}(x) = W_{p,q}(\tau_{k,l}x)$  by the invariance property of  $W_{p,q}$ . This shows that also  $\tau_{k,l}x$  is a minimizer, whence, by the previous corollary, either  $\tau_{k,l}x \ll x$ ,  $\tau_{k,l}x = x$  or  $\tau_{k,l}x \gg x$ . In particular,  $x$  is a Birkhoff configuration.  $\square$

**Lemma 4.7.** *Let  $n \in \mathbb{N}$ . Every  $p, q$ -minimizer is an  $np, nq$ -minimizer and vice versa.*

**Proof.** Assume that  $x$  is an  $np, nq$ -minimizer, that is a minimizer of  $W_{np, nq}$  on  $\mathbb{X}_{np, nq}$ . Then, by Aubry's lemma,  $x \in \mathcal{B}_{np, nq}$ . Theorem 3.12 now implies that  $\mathcal{B}_{np, nq} = \mathcal{B}_{p, q}$ , so actually  $x \in \mathcal{B}_{p, q} \subset \mathbb{X}_{p, q}$ . Note now that on  $\mathbb{X}_{p, q}$  it holds that  $W_{np, nq} = n^d W_{p, q}$  and let  $y \in \mathbb{X}_{p, q} \subset \mathbb{X}_{np, nq}$ . Then  $W_{p, q}(x) = n^{-d} W_{np, nq}(x) \leq n^{-d} W_{np, nq}(y) = W_{p, q}(y)$ . Thus,  $x$  is a  $p, q$ -minimizer.

In the other direction, if  $x$  is a  $p, q$ -minimizer and  $y$  is an  $np, nq$ -minimizer, then  $y \in \mathcal{B}_{p, q}$  and  $W_{np, nq}(x) = n^d W_{p, q}(x) \leq n^d W_{p, q}(y) = W_{np, nq}(y)$ , that is  $x$  is an  $np, nq$ -minimizer.  $\square$

The following result shows that  $p, q$ -minimizers are global minimizers. Recall that  $x$  is called a global minimizer if for every finite set  $B \subset \mathbb{Z}^d$  and every  $y$  with support in its  $r$ -interior  $\hat{B}^{(r)}$ , one has that  $W_B(x + y) \geq W_B(x)$ , with  $W_B(x) := \sum_{j \in B} S_j(x)$ .

**Theorem 4.8.** *Periodic minimizers are global minimizers.*

**Proof.** Let  $x \in \mathbb{X}_{p, q}$  be a  $p, q$ -minimizer. If  $x$  is not a global minimizer, then there exists a finite set  $B \subset \mathbb{Z}^d$  and a configuration  $y$  with  $\text{supp}(y) \subset \hat{B}^{(r)}$ , such that  $W_B(x + y) < W_B(x)$ . Since  $B$  is finite, there exist a  $k \in \mathbb{Z}^d$  and an  $n \in \mathbb{N}$  such that  $\text{supp}(y) \subset B \subset k + B_{np}$ . Now define  $\tilde{y} \in \mathbb{X}_{np, nq}$  by setting  $\tilde{y}_i = y_j$  when  $j$  is the unique point in  $k + B_{np}$  for which  $j = i \bmod np$  ( $\mathbb{Z}^d$ ). In other words,  $\tilde{y}$  is the  $np$ -periodic extension of  $y|_{k+B_{np}}$ . Then we conclude that

$$W_{np, nq}(x + \tilde{y}) - W_{np, nq}(x) = W_{k+nB_p}(x + y) - W_{k+nB_p}(x) = W_B(x + y) - W_B(x) < 0,$$

so  $x$  is not  $np, nq$ -minimizer. This contradicts Lemma 4.7.  $\square$

Perhaps surprisingly, to prove the converse one needs to be slightly more ingenious. We have not found this statement anywhere in the literature:

**Theorem 4.9.** *If  $x \in \mathbb{X}_{p, q}$  is a global minimizer, then it is a  $p, q$ -minimizer.*

**Proof.** Suppose that  $x \in \mathbb{X}_{p, q}$  is not a  $p, q$ -minimizer. We will prove that this implies that  $x$  is not a global minimizer. Our assumption means that there is a  $y \in \mathbb{X}_{p, q}$  for which  $0 < \varepsilon := W_{p, q}(x) - W_{p, q}(y)$ . This in turn implies that  $W_{np, nq}(x) - W_{np, nq}(y) = n^d \varepsilon$ .

By periodicity, we may assume that  $x \ll y$ . Let us now define, for  $n \in \mathbb{N}$ , the configurations  $x \leq y^n \leq y$  by

$$y_i^n = \begin{cases} y_i & \text{if } i \in \hat{B}_{np}^{(r)}, \\ x_i & \text{otherwise.} \end{cases}$$

Here,  $\hat{B}_{np}^{(r)}$  is the  $r$ -interior of  $B_{np}$ . By definition,  $y^n$  is a variation of  $x$  with support in this  $r$ -interior. It now holds that

$$\begin{aligned} W_{B_{np}}(x) - W_{B_{np}}(y^n) &= \sum_{j \in B_{np}} (S_j(x) - S_j(y^n)) = \sum_{j \in B_{np}} (S_j(x) - S_j(y)) + (S_j(y) - S_j(y^n)) \\ &= n^d \varepsilon + \sum_{j \in B_{np}} (S_j(y) - S_j(y^n)). \end{aligned}$$

Because the support of  $y - y^n$  is contained in  $\mathbb{Z}^d \setminus \hat{B}_{np}^{(r)}$  and the range of interaction of the  $S_j$  is equal to  $r$ , the number of nonzero terms in the above sum is at most  $(2r)^{2d+1} |\partial B_{np}| \leq E n^{d-1}$ , where  $E$  is a constant depending only on  $r, d$  and  $p$ .

Moreover, by compactness of  $[x, y] := \{z \mid x \leq z \leq y\}$ , there is a constant  $e > 0$  so that  $|S_j(y)|, |S_j(y^n)| < e$ . This then implies that

$$W_{B_{np}}(x) - W_{B_{np}}(y^n) \geq \varepsilon n^d - 2Ee \cdot n^{d-1}.$$

Choosing  $n$  large enough, we see that  $x$  is not a global minimizer.  $\square$

#### 4.2. Nonperiodic minimizers

In this section, we show that global minimizers of all rotation vectors exist. They are constructed as limits of periodic minimizers. Moreover, we show that they satisfy a certain pairwise regularity. The results in this section are standard.

**Lemma 4.10.** *The set of global minimizers is closed in the topology of pointwise convergence.*

**Proof.** Assume that  $x^n$  is a sequence of global minimizers converging pointwise to  $x^\infty$ . Let  $B \subset \mathbb{Z}^d$  be a finite set and  $y$  a configuration with support in  $\dot{B}^{(r)}$ . Then

$$W_B(x^n + y) - W_B(x^n) \geq 0. \quad (4.11)$$

But  $W_B$  is continuous with respect to pointwise convergence, so that taking the limit for  $n \rightarrow \infty$  of Eq. (4.11), we find that  $W_B(x^\infty + y) - W_B(x^\infty) \geq 0$ . So  $x^\infty$  is a global minimizer.  $\square$

**Theorem 4.11.** *For all rotation vectors  $\omega \in \mathbb{R}^d$  and all local potentials  $S_j$ , there exists a global minimizer  $x \in \bar{\mathcal{B}}_\omega$ .*

**Proof.** For any  $\omega \in \mathbb{R}^d$ , we can take a sequence  $\omega_n \in \mathbb{Q}^d$ , such that  $\lim_{n \rightarrow \infty} \omega_n = \omega$ , while  $\langle \omega_n, k \rangle + l = 0$  for all the  $k$  and  $l$  for which  $\langle \omega, k \rangle + l = 0$ . We take a corresponding sequence  $(p_n, q_n)$  of principal periods for which  $\omega_n := -p_n^{-T} q_n$ . By Theorems 4.3 and 4.8, there exists a global minimizer  $x^n \in \mathcal{B}_{p_n, q_n} = \bar{\mathcal{B}}_{\omega_n}$ . In particular,  $x_n$  has rotation vector  $\omega_n$  and satisfies  $\tau_{k,l} x^n = x^n$  for all  $k$  and  $l$  for which  $\langle \omega, k \rangle + l = 0$ . Because the  $\omega_n$  and  $\omega$  lie in some compact subset  $K$  of  $\mathbb{R}^d$ , Proposition 3.6, guarantees that there is a subsequence of the  $x^n$  that converges pointwise to a Birkhoff configuration  $x^\infty \in \mathcal{B}_K$ . By continuity of the rotation vector  $x \mapsto \omega(x)$ , see Proposition 3.5,  $x^\infty$  actually has rotation vector  $\omega$ . Moreover, the limit  $x^\infty$  will have the same periodicities: denoting the converging subsequence also by  $x^n$ , the continuity of  $\tau_{k,l}$  implies that  $\tau_{k,l} x^\infty = \tau_{k,l}(\lim_{n \rightarrow \infty} x^n) = \lim_{n \rightarrow \infty} \tau_{k,l} x^n = \lim_{n \rightarrow \infty} x^n = x^\infty$  for all  $k$  and  $l$  with  $\langle \omega, k \rangle + l = 0$ . Finally,  $x^\infty$  is a global minimizer by Theorem 4.8 and Lemma 4.10.  $\square$

The following result expresses the regularity of pairwise comparable stationary solutions. It is the analogue of a Harnack inequality for elliptic PDEs.

**Theorem 4.12 (Elliptic Harnack inequality).** *Let  $x < y$  be two Birkhoff configurations with rotation vector in the compact set  $K \subset \mathbb{R}^d$ . Suppose that  $x$  and  $y$  are stationary for the local potentials  $S_j$ . Then there is a constant  $\delta$ , depending only on  $K$  and  $\|i - k\|$ , such that for all  $i$  and  $k$ ,*

$$(y_k - x_k) \leq \delta(y_i - x_i).$$

*In particular, if  $x < y$ , then  $x \ll y$ .*

**Proof.** By interpolation: let  $x$  and  $y$  be stationary and Birkhoff and let  $i, k \in \mathbb{Z}^d$  and assume first that  $\|i - k\| = 1$ . Choose a  $B$  with  $i \in \dot{B}^{(r)}$  and recall the definition  $W_B(x) = \sum_{j \in B} S_j(x)$ . Then, by

stationarity,

$$\begin{aligned} 0 &= \partial_i W_B(y) - \partial_i W_B(x) = \int_0^1 \frac{d}{d\tau} \left( \sum_{\|i-j\| \leq r} \partial_i S_j(\tau y + (1-\tau)x) \right) d\tau \\ &= \sum_{\|i-j\| \leq r, \|j-l\| \leq r} \left( \int_0^1 \partial_{i,l} S_j(\tau y + (1-\tau)x) d\tau \right) (y_l - x_l). \end{aligned}$$

Since, by the twist condition C, the only possibly positive terms on the right-hand side are the  $(\partial_{i,i} S_j(\tau y + (1-\tau)x))(y_i - x_i)$ , the right-hand side is less than or equal to

$$\begin{aligned} &\left( \int_0^1 \sum_{\|i-j\| \leq r} \partial_{i,i} S_j(\tau y + (1-\tau)x) d\tau \right) (y_i - x_i) \\ &+ \left( \int_0^1 \sum_{\|k-i\|=1} \partial_{i,k} S_i(\tau y + (1-\tau)x) d\tau \right) (y_k - x_k). \end{aligned}$$

Now, because  $x$  and  $y$  are Birkhoff, so is every  $\tau y + (1-\tau)x$  and hence by Corollary 3.7, there are constants  $\lambda, C > 0$ , depending only on the compact set  $K$ , such that for all  $j$  and all  $\|i-k\|=1$ , it holds that  $\partial_{i,k} S_i < -\lambda$ , while  $\partial_{i,i} S_j < C$  for all  $i$  and  $j$ . Thus,

$$0 \leq (2r)^d C (y_i - x_i) - 2d\lambda (y_k - x_k).$$

This proves the theorem for  $\|i-k\|=1$  with  $\delta = \delta_1 := (2r)^d C / 2d\lambda$ . For  $\|i-k\| > 1$ , the result then follows by induction and it holds for  $\delta = \delta_1^{\|i-k\|}$ .  $\square$

#### 4.3. Aubry–Mather sets

We make the following definition:

**Definition 4.13.** An Aubry–Mather set  $\mathcal{M} \subset \mathbb{R}^{\mathbb{Z}^d}$  is a collection of configurations with the following properties:

- $\mathcal{M}$  is nonempty and closed under pointwise convergence.
- $\mathcal{M}$  is strictly ordered, i.e. for every  $x, y \in \mathcal{M}$ ,  $x \ll y$ ,  $x = y$  or  $x \gg y$ .
- $\mathcal{M}$  is shift-invariant: if  $x \in \mathcal{M}$ , then for every  $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}$ , also  $\tau_{k,l} x \in \mathcal{M}$ .
- Every  $x \in \mathcal{M}$  is a global minimizer of the variational recurrence relation (2.8).
- $\mathcal{M}$  does not contain any strictly smaller set with the properties listed above.

The strict ordering and the shift-invariance of an Aubry–Mather set  $\mathcal{M}$  imply that any configuration  $x \in \mathcal{M}$  is Birkhoff and hence has a rotation vector  $\omega = \omega(x)$ . The ordering of  $\mathcal{M}$  moreover implies that this rotation vector is independent of the choice of  $x \in \mathcal{M}$ , that is  $\omega = \omega(\mathcal{M})$  and thus,  $\mathcal{M} \subset \mathcal{B}_\omega$ .

Recall that Theorem 4.11 states that for every rotation vector  $\omega$  there exists a minimizer  $x \in \mathcal{B}_\omega$  for which  $\tau_{k,l} x = x$  as soon as  $\langle \omega, k \rangle + l = 0$ . This in fact implies that a certain Aubry–Mather



set  $\mathcal{M}(x) \subset \mathcal{B}_\omega$  exists. This  $\mathcal{M}(x)$  is constructed as follows. One starts by defining the collection  $\tilde{\mathcal{M}}(x) \subset \mathcal{B}_\omega$  as the closure with respect to pointwise convergence of the set of translates of  $x$ :

$$\tilde{\mathcal{M}}(x) := \overline{\{\tau_{k,l}x \mid (k, l) \in \mathbb{Z}^d \times \mathbb{Z}\}}.$$

This is almost an Aubry–Mather set:

**Lemma 4.14.** *Let  $x \in \mathcal{B}_\omega$  be an action-minimizer with the property that  $\tau_{k,l}x = x$  when  $\langle \omega, k \rangle + l = 0$ . Then  $\tilde{\mathcal{M}}(x)$  is nonempty, closed, strictly ordered, shift-invariant and consists of minimizers. Moreover, for every  $y \in \tilde{\mathcal{M}}(x)$  it holds that  $\tau_{k,l}y = y$  as soon as  $\langle \omega, k \rangle + l = 0$ . When  $\omega \in \mathbb{Q}^d$ , then  $\tilde{\mathcal{M}}(x)$  is an Aubry–Mather set.*

**Proof.** By definition,  $\tilde{\mathcal{M}}(x)$  is nonempty and closed.

We note that when  $x$  is a minimizer, then so is  $\tau_{k,l}x$  and because any pointwise limit of minimizers is a minimizer itself, by Lemma 4.10, we see that  $\tilde{\mathcal{M}}(x)$  consists of minimizers only.

When  $y \in \tilde{\mathcal{M}}(x)$ , say  $y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}x$ , then the continuity of  $\tau_{k,l}$  implies that  $\tau_{k,l}y = \tau_{k,l}(\lim_{n \rightarrow \infty} \tau_{k_n, l_n}x) = \lim_{n \rightarrow \infty} \tau_{k,l}(\tau_{k_n, l_n}x) = \lim_{n \rightarrow \infty} \tau_{k+k_n, l+l_n}x$  and thus,  $\tilde{\mathcal{M}}(x)$  is shift-invariant.

The fact that  $x$  is a Birkhoff configuration means that the collection  $\{\tau_{k,l}x \mid (k, l) \in \mathbb{Z}^d \times \mathbb{Z}\}$  is ordered. Now let  $y$  and  $z$  be elements of  $\tilde{\mathcal{M}}(x)$ , say  $y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}x$  and  $z = \lim_{n \rightarrow \infty} \tau_{K_n, L_n}x$ . We claim that  $y \leq z$  or  $z \leq y$ . If not, then there are  $i, k$  with  $y_i < z_i$  and  $y_k > z_k$ . The pointwise convergence then implies that there are  $n$  and  $N$  so that  $(\tau_{k_n, l_n}x)_i < (\tau_{K_N, L_N}x)_i$  and  $(\tau_{k_n, l_n}x)_k > (\tau_{K_N, L_N}x)_k$ . This is a contradiction. The second conclusion of Theorem 4.12 now implies that  $y \ll z$ ,  $y = z$  or  $y \gg z$ , that is  $\tilde{\mathcal{M}}(x)$  is strictly ordered.

The penultimate conclusion of the lemma follows from the continuity of  $\tau_{k,l}$  and the fact that  $\tau_{k,l}x = x$  when  $\langle \omega, k \rangle + l = 0$ . Namely, for such  $k$  and  $l$  and for  $y \in \tilde{\mathcal{M}}(x)$ , say  $y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}x$ , we have that  $\tau_{k,l}y = \tau_{k,l}(\lim_{n \rightarrow \infty} \tau_{k_n, l_n}x) = \lim_{n \rightarrow \infty} \tau_{k,l}(\tau_{k_n, l_n}x) = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}(\tau_{k,l}x) = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}x = y$ .

Finally, when  $\omega \in \mathbb{Q}^d$ , then our assumptions imply that  $x$  is periodic, say  $x \in \mathbb{X}_{p,q}$ , with  $(p, q)$  a collection of principal periods for  $\omega$ . This implies that the  $\tau$ -orbit of  $x$  is finite. Thus,  $\tilde{\mathcal{M}}(x)$  is equal to this single  $\tau$ -orbit and cannot contain any proper nonempty  $\tau$ -invariant subset.  $\square$

It is clear from the proof of Lemma 4.14, that when  $\omega \in \mathbb{Q}^d$ , then every Aubry–Mather set is finite and consists of the translates of one periodic minimizer. Thus, the Aubry–Mather sets of rational rotation vector do not need to be unique.

On the other hand, when  $\omega \in \mathbb{R}^d \setminus \mathbb{Q}^d$  is irrational, then  $\tilde{\mathcal{M}}(x)$  may fail to be an Aubry–Mather set. Then one replaces  $\tilde{\mathcal{M}}(x)$  by its recurrent subset

$$\mathcal{M}(x) := \left\{ y \in \tilde{\mathcal{M}}(x) \mid y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}y \text{ for a sequence } (k_n, l_n) \text{ with } \langle \omega, k_n \rangle + l_n \neq 0 \right\}.$$

Before proving that this  $\mathcal{M}(x)$  is indeed an Aubry–Mather set, let us define for a configuration  $y \in \tilde{\mathcal{M}}(x)$ , the configurations

$$y^- := \sup\{\tau_{k,l}y \ll y\} \quad \text{and} \quad y^+ := \inf\{\tau_{k,l}y \gg y\}.$$

We remark that, by definition,  $y \in \mathcal{M}(x)$  if and only if  $y = y^-$  or  $y = y^+$ , or both. We now have the following technical result:

**Proposition 4.15.** *For  $y, z \in \tilde{\mathcal{M}}(x)$  it holds that  $y^- = \sup\{\tau_{k,l}z \ll y^-\}$  and  $y^+ = \inf\{\tau_{k,l}z \gg y^+\}$ .*

**Proof.** Let us prove the first equality: the proof of the second one is similar. We denote  $z^-(y^-) := \sup\{\tau_{k,l}z \ll y^-\}$  and we argue by contradiction. That is, we suppose that  $z^-(y^-) \neq y^-$ , and hence,

that  $z^-(y) \ll y^-$ . Then, because  $y^-$  can be approximated from below by translates of  $y$  by definition, there are  $k$  and  $l$  so that  $z^-(y^-) \ll \tau_{k,l}y \ll y^-$ . This implies that  $\tau_{k,l}y \ll y$  and in view of Proposition 3.8, we must therefore have that  $\langle \omega, k \rangle + l < 0$ . Applying  $\tau_{-k,-l}$  to the inequality  $z^-(y^-) \ll \tau_{k,l}y$ , we obtain that  $\tau_{-k,-l}z^-(y^-) \ll y$ . But because  $\langle \omega, -k \rangle - l > 0$ , we must also have that  $z^-(y^-) \ll \tau_{-k,-l}z^-(y^-)$ . Hence,  $z^-(y^-) \ll \tau_{-k,-l}z^-(y^-) \ll y$ . But this contradicts the definition of  $z^-(y^-)$ , because by continuity of  $\tau_{-k,-l}$ , if  $z^-(y^-) = \lim_{n \rightarrow \infty} \tau_{k_n, l_n} z$ , then also  $\tau_{-k,-l}z^-(y^-) = \lim_{n \rightarrow \infty} \tau_{-k+k_n, -l+l_n} z$  is a limit of translates of  $z$  that lie below  $y^-$ .  $\square$

We are now ready to prove:

**Theorem 4.16.** *When  $\omega \in \mathbb{R}^d / \mathbb{Q}^d$ , then  $\mathcal{M}(x)$  is the unique Aubry–Mather set contained in  $\tilde{\mathcal{M}}(x)$ .*

**Proof.** Proposition 4.15 says that any  $y \in \mathcal{M}(x)$  is a limit point of the  $\tau$ -orbit of any  $z \in \tilde{\mathcal{M}}(x)$ . Thus, any nonempty, shift-invariant closed subset of  $\tilde{\mathcal{M}}(x)$  should contain  $\mathcal{M}(x)$ . It remains to show that  $\mathcal{M}(x)$  is nonempty, shift-invariant and closed.

First of all, Proposition 4.15 applied to  $z = y^-$  and  $z = y^+$  respectively, says that  $(y^-)^- = y^-$  and  $(y^+)^+ = y^+$ , i.e. that  $y^-$  and  $y^+$  are recurrent. This shows that  $\mathcal{M}(x)$  is nonempty.

Shift-invariance of  $\mathcal{M}(x)$  follows from the continuity of  $\tau_{k,l}$ : when  $y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n} y$ , then  $\tau_{k,l}y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n}(\tau_{k,l}y)$ .

To prove that  $\mathcal{M}(x)$  is closed, assume that  $\lim_{n \rightarrow \infty} y_n = y$  pointwise for a sequence  $y_n$  of recurrent configurations. When the limit  $y$  is not recurrent, then  $y^- \ll y \ll y^+$ , so that there is an  $n$  for which  $y^- \ll y_n \ll y^+$ . But  $y_n$  is recurrent, hence  $y_n \neq y$ , while by Proposition 4.15,  $y_n$  can be approximated by translates of  $y$ . Hence, there are  $k$  and  $l$  such that  $y^- \ll \tau_{k,l}y \ll y^+$  and  $\tau_{k,l}y \neq y$ . This contradicts the definition of  $y^-$  or  $y^+$ .  $\square$

**Remark 4.17.** A theorem of Bangert [2] in the case of elliptic PDEs, states that when  $\omega \in \mathbb{R}^d \setminus \mathbb{Q}^d$ , then the recurrent subset actually does not depend on the choice of the Birkhoff minimizer  $x \in \mathcal{B}_\omega$ . In other words, that when  $x, y \in \mathcal{B}_\omega$  are such that  $\tau_{k,l}x = x$  and  $\tau_{k,l}y = y$  whenever  $\langle \omega, k \rangle + l = 0$ , then  $\mathcal{M}(x) = \mathcal{M}(y)$ .

The proof of this theorem is nontrivial. The essence of it lies in proving an Aubry lemma for recurrent minimizers, that is to show that if  $\tilde{x} \in \mathcal{M}(x)$  and  $\tilde{y} \in \mathcal{M}(y)$  are recurrent, then  $\tilde{x} \ll \tilde{y}$ ,  $\tilde{x} = \tilde{y}$  or  $\tilde{x} \gg \tilde{y}$ .

We claim that a similar theorem holds for lattices instead of PDEs, but we will not prove this, as it is not essential for the remainder of this paper. As a result, the Aubry–Mather set of an irrational rotation vector is unique.

The following well-known result shows that the set of recurrent minimizers can have a complicated topology. We recall that a topological space  $\mathcal{C}$  is called a *Cantor set* if it is closed, perfect and totally disconnected. “Perfect” means that every element  $c \in \mathcal{C}$  is a limit of points in  $\mathcal{C} \setminus \{c\}$ . “Totally disconnected” means that for any two elements  $c_1, c_2 \in \mathcal{C}$  one can decompose  $\mathcal{C}$  as the disjoint union of closed sets  $\mathcal{C}_1$  and  $\mathcal{C}_2$  with  $c_1 \in \mathcal{C}_1$  and  $c_2 \in \mathcal{C}_2$ .

**Theorem 4.18.** *If  $\omega \in \mathbb{R}^d \setminus \mathbb{Q}^d$ , then  $\mathcal{M}(x)$  is either connected or a Cantor set.*

**Proof.** The recurrent subset is perfect by definition: for every  $y \in \mathcal{M}(x)$ , it holds that  $y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n} y$ , where by Proposition 3.8 the condition that  $\langle \omega, k_n \rangle + l_n \neq 0$  guarantees that  $\tau_{k_n, l_n} y \neq y$  for all  $n$ .

We will now show that when  $\mathcal{M}(x)$  is not connected, then there is a  $y \in \mathcal{M}(x)$  so that  $y^- \neq y^+$ . So let's assume that  $\mathcal{M}(x)$  is not connected and write  $\mathcal{M}(x) = U \cup V$  for two nonempty closed subsets  $U$  and  $V$  with  $U \cap V = \emptyset$ . We may assume that there exist  $u \in U$  and  $v \in V$  so that  $u \ll v$ ,

whence we can define  $y := \sup\{u \in U \mid u \ll v\}$ . Clearly,  $y \in U$ , because  $U$  is closed. Hence,  $y \ll v$ . We claim that  $y^+ \neq y$ . This is easily proved: if  $y^+ = y$ , then  $y = \lim_{n \rightarrow \infty} \tau_{k_n, l_n} y$  for a sequence with  $y \ll \tau_{k_n, l_n} y \ll v$ . By definition of  $y$ , it must hold that  $\tau_{k_n, l_n} y \in V$ . Hence, because  $V$  is closed, also  $y \in V$ , which is a contradiction.

The next step is to observe that an order interval  $[y^-, y^+] := \{z \in \mathbb{R}^{\mathbb{Z}^d} \mid y^- \leq z \leq y^+\}$  can never contain any recurrent elements other than  $y^-$  and  $y^+$ . Namely, if  $y^- \ll v \ll y^+$  were such a recurrent element, then by Proposition 4.15, it can be approximated by translates of  $y$ , so that there are  $k$  and  $l$  with  $y^- \ll \tau_{k, l} y \ll y^+$ . This contradicts the definition of  $y^-$  or  $y^+$ . This is why we call the order interval  $[y^-, y^+]$  a *gap* in the Aubry–Mather set.

Now we show that when  $\mathcal{M}(x)$  is not connected, and hence contains at least one gap  $[y^-, y^+]$ , then between any two elements  $w, z \in \mathcal{M}(x)$  there exists a gap. Namely, for any given pair  $w \ll z$ , either  $[w, z]$  is a gap, or there is a recurrent element  $w \ll u \ll z$ . By Proposition 4.15, this  $u$  can then be approximated by the  $\tau$ -orbit of  $y^-$ , which implies that there are  $k$  and  $l$  so that  $w \ll \tau_{k, l} y^- \ll z$ . But when  $[y^-, y^+]$  is a gap, then so is  $[\tau_{k, l} y^-, \tau_{k, l} y^+]$ , since  $\tau_{k, l}$  is order-preserving. We must therefore have that  $w \ll \tau_{k, l} y^- \ll \tau_{k, l} y^+ \leq z$ , i.e. that there is a gap between  $w$  and  $z$ .

This implies that  $\mathcal{M}(x)$  is totally disconnected: if  $w, z \in \mathcal{M}(x)$  with  $w \ll z$ , then there is a gap  $[y^-, y^+]$  with  $w \leq y^- \ll y^+ \leq z$  and hence  $\mathcal{M}(x)$  splits as the disjoint union of the closed sets  $\{u \in \mathcal{M}(x) \mid u \leq y^-\}$  and  $\{v \in \mathcal{M}(x) \mid v \geq y^+\}$  that contain  $w$  and  $z$  respectively.  $\square$

The proof of Theorem 4.18 shows that for any  $y \in \tilde{\mathcal{M}}(x)$ , in the order interval

$$[y^-, y^+] := \{z \in \mathbb{R}^{\mathbb{Z}^d} \mid y^- \leq z \leq y^+\}$$

only the elements  $y^-$  and  $y^+$  are recurrent. Hence, when  $y^- \neq y^+$ , then  $[y^-, y^+]$  is called a *gap* in the Aubry–Mather set. Moreover, in the case that  $\mathcal{M}(x)$  is not connected, then between any two recurrent configurations there exists such a gap.

When  $\mathcal{M}(x)$  is connected, then we say that it forms a *foliation*: for every  $i \in \mathbb{Z}^d$  and every  $\xi \in \mathbb{R}$  there is a unique  $y \in \mathcal{M}(x)$  so that  $y_i = \xi$ . In the case that  $\mathcal{M}(x)$  is a Cantor set, one says that it forms a *lamination*: for every  $i$  and every  $\xi$  there is at most one  $y$  so that  $y_i = \xi$ .

Both foliations and laminations by minimizers occur in examples, for instance that of the Frenkel–Kontorova lattice (1.1). In fact, when  $V(\xi) \equiv 0$ , then the Aubry–Mather sets are all of the form  $\mathcal{M}(x^{\omega, 0}) := \{x^{\omega, \xi} \mid \xi \in \mathbb{R}\}$ , where we recall that the linear configuration  $x^{\omega, \xi}$  is defined by  $x_i^{\omega, \xi} = \langle \omega, i \rangle + \xi$ . These Aubry–Mather sets are clearly connected.

On the other hand, the following theorem says that when the onsite potentials  $V(\xi)$  are sufficiently oscillatory, then the Aubry–Mather sets must be Cantor sets:

**Theorem 4.19.** *Let  $S_j$  be local potentials satisfying conditions A–E and let  $K \subset \mathbb{R}^d$  be a compact set. Then there exists a number  $M > 0$ , depending on the  $S_j$  and on  $K$ , such that for every 1-periodic twice continuously differentiable function  $V = V(\xi)$  with  $\text{osc } V := \max_{\xi, v \in \mathbb{R}} (V(\xi) - V(v)) > M$ , the collection of local potentials  $\tilde{S}_j$  defined by  $\tilde{S}_j(x) = S_j(x) + V(x_j)$  does not possess any connected, strictly ordered shift-invariant family of global minimizers of rotation vector  $\omega \in K$ .*

**Proof.** Because  $\mathcal{B}_K/\mathbb{Z}$  is compact and the functions  $S_j$  are  $\tau$ -invariant and continuous, their oscillation over  $\mathcal{B}_K$  is bounded and, say, equal to  $N := \text{osc}_{\mathcal{B}_K} S_j = \max_{x, y \in \mathcal{B}_K} (S_j(x) - S_j(y))$ . Let  $M > (2r + 1)^d N$ , where  $r \geq 1$  is the finite interaction range of the local potentials  $S_j$ , and choose a smooth 1-periodic onsite potential  $V$  with oscillation larger than  $M$ . Assume for instance that  $V(\xi) - V(v) > M$  for certain  $\xi, v \in \mathbb{R}$ .

We will now prove that if a configuration  $x \in \mathcal{B}_K$  has  $x_0 = \xi$ , then it cannot be a global minimizer. In other words, that  $x$  is a “gap configuration”. This is easily shown by defining  $y : \mathbb{Z}^d \rightarrow \mathbb{R}$  by setting  $y_i = 0$  for  $i \neq 0$  and  $y_0 = v - \xi$ . Now choose a finite subset  $B \subset \mathbb{Z}^d$  such that  $B_0^r \subset B$ . Then  $\text{supp } y = \{0\} \subset \dot{B}^{(r)}$  and we compute that

$$\begin{aligned}\tilde{W}_B(x) - \tilde{W}_B(x+y) &= \sum_{j \in B} \tilde{S}_j(x) - \tilde{S}_j(x+y) = \sum_{j \in B_0^r} \tilde{S}_j(x) - \tilde{S}_j(x+y) \\ &= V(\xi) - V(v) + \sum_{j \in B_0^r} S_j(x+y) - S_j(x) > M - (2r+1)^d N > 0.\end{aligned}$$

This shows that  $x$  is not a global minimizer.  $\square$

**Example 4.20.** For the Frenkel–Kontorova lattice, Theorem 4.19 can be improved upon considerably. In fact, by Lemma 3.5,  $\text{osc}_{\mathcal{B}_\omega}(x_j - x_k) \leq 2$ , which is independent of  $\omega$ . Therefore, the oscillation over  $\mathcal{B}$  of the interaction potential  $\frac{1}{8d} \sum_{\|j-k\|=1} (x_j - x_k)^2$  is bounded above by 1. Thus, for any on-site potential  $V(\xi)$  with oscillation larger than  $2d$ , the Frenkel–Kontorova lattice with local potentials  $S_j(x) = \frac{1}{8d} \sum_{\|k-j\|=1} (x_j - x_k)^2 + V(x_j)$  does not have a connected family of global minimizers of any rotation vector at all.

The latter result for the Frenkel–Kontorova lattice is well known in dimension  $d = 1$ . It turns out that the one-dimensional Frenkel–Kontorova lattice is equivalent to the Chirikov standard map  $T_V$ , see Appendix A. As such, Theorem 4.19 and the discussion above say that for any on-site potential  $V$  with oscillation larger than 2, the standard map  $T_V$  has no rotational invariant curves. In the case that  $V$  has the “standard” form  $V(\xi) = \frac{k}{8\pi^2} \cos(2\pi\xi)$ , so that  $\text{osc } V = \frac{k}{4\pi^2}$ , we obtain that there are no rotational invariant curves for  $k > 8\pi^2$ . In fact, in this case the much stronger computer-proved bound  $k > \frac{63}{64}$  is actually known, see [12].

## 5. A formal gradient flow

The idea of studying globally stationary solutions by means of a formal gradient flow goes back to Golé, see [7]. We will review his ideas in this section. The new result is a parabolic Harnack inequality, see Theorem 6.4.

The study of the formal gradient flow starts with the observation that one can assign a meaning to the partial derivatives of the formal, and generally nonconvergent sum  $W(x) = \sum_{j \in \mathbb{Z}^d} S_j(x)$ , namely as follows. Since the potentials  $S_j$  are of finite range, every variable  $x_i$  appears only in finitely many terms of the formal series. Hence, we may write, with a slight abuse of notation,

$$(\nabla W(x))_i := \partial_i W(x) = \sum_{\|j-i\| \leq r} \partial_i S_j(x).$$

Note that  $\nabla W : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$  is well defined as soon as the  $S_j$  are continuously differentiable and that  $\nabla W$  is the formal gradient of  $W$  with respect to the  $l_2$ -inner product  $\langle x, y \rangle_2 = \sum_{j \in \mathbb{Z}^d} x_j y_j$ .

We remark that  $x$  is globally stationary if and only if  $\nabla W(x) = 0$ . In this section, we shall nevertheless view such  $x$  as stationary points of the auxiliary differential equation

$$\frac{dx}{dt} = -\nabla W(x).$$

This differential equation shall be defined on an appropriate Banach subspace  $\mathbb{X} \subset \mathbb{R}^{\mathbb{Z}^d}$  of configurations, for which its initial value problem has existence and uniqueness of solutions. The corresponding flow is called the negative gradient flow of  $W$ . The motivation to study the negative gradient flow is simply that it will help us find globally stationary solutions.

The Banach subspace we choose to work with is the exponentially weighted configuration space

$$\mathbb{X} := \left\{ x \in \mathbb{R}^{\mathbb{Z}^d} \mid \|x\|_{\mathbb{X}} := \sum_{i \in \mathbb{Z}^d} \frac{|x_i|}{2^{\|i\|}} < \infty \right\} \subset \mathbb{R}^{\mathbb{Z}^d},$$

where we recall that  $\|i\| := \sum_{k=1}^d |i_k|$ . First of all, the space of Birkhoff configurations is contained in  $\mathbb{X}$ :

**Lemma 5.1.**  $\mathcal{B} \subset \mathbb{X}$ .

**Proof.** This follows because every  $x \in \mathcal{B}$  has a rotation vector, say  $\omega$ , and  $|x_i - x_0 - \langle \omega, i \rangle| \leq 1$ . This implies that  $|x_i| \leq \|\omega\| \cdot \|i\| + |x_0| + 1$  and hence

$$\|x\|_{\mathbb{X}} = \sum_{i \in \mathbb{Z}^d} \frac{|x_i|}{2^{\|i\|}} \leq \sum_{i \in \mathbb{Z}^d} \frac{\|i\| \|\omega\| + |x_0| + 1}{2^{\|i\|}} < \infty. \quad \square$$

We moreover note that the topology  $\mathcal{B}$  inherits from  $\mathbb{X}$  is exactly that of pointwise convergence:

**Proposition 5.2.** Let  $x \in \mathbb{X}$  and for all  $n \in \mathbb{N}$ , let  $x_n \in \mathbb{X}$ . Then  $\lim_{n \rightarrow \infty} \|x_n - x\|_{\mathbb{X}} = 0$  if and only if  $\lim_{n \rightarrow \infty} x_n = x$  pointwise. In particular, a sequence in  $\mathcal{B}$  converges in  $\mathbb{X}$  if and only if it converges pointwise.

**Proof.** The first claim is obvious. The second claim follows because  $\mathcal{B}$  is a closed subset of  $\mathbb{X}$ .  $\square$

Before showing the existence of the negative gradient flow on  $\mathbb{X}$ , we need the following simple lemma, which shows that the shift maps  $\tau_{k,l}$  are Lipschitz on  $\mathbb{X}$ :

**Lemma 5.3.** Let  $x, y \in \mathbb{X}$  and  $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}$ . Then  $\tau_{k,l}x \in \mathbb{X}$  and  $\tau_{k,l}y \in \mathbb{X}$ , while  $\|\tau_{k,l}x - \tau_{k,l}y\|_{\mathbb{X}} \leq 2^{\|k\|} \|x - y\|_{\mathbb{X}}$ .

**Proof.** First of all,

$$\|\tau_{k,0}x\|_{\mathbb{X}} = \sum_{i \in \mathbb{Z}^d} \frac{|x_{i+k}|}{2^{\|i\|}} = 2^{\|k\|} \sum_{i \in \mathbb{Z}^d} \frac{|x_{i+k}|}{2^{\|i\|+\|k\|}} \leq 2^{\|k\|} \sum_{i \in \mathbb{Z}^d} \frac{|x_{i+k}|}{2^{\|i+k\|}} = 2^{\|k\|} \|x\|_{\mathbb{X}}.$$

Therefore,  $\|\tau_{k,l}x\|_{\mathbb{X}} = \|\tau_{k,0}x + l\|_{\mathbb{X}} \leq \|\tau_{k,0}x\|_{\mathbb{X}} + \|l\|_{\mathbb{X}} \leq 2^{\|k\|} \|x\|_{\mathbb{X}} + \|l\|_{\mathbb{X}} < \infty$  and similarly for  $\tau_{k,l}y$ . In particular,  $\|\tau_{k,l}x - \tau_{k,l}y\|_{\mathbb{X}} = \|\tau_{k,0}(x - y)\|_{\mathbb{X}} \leq 2^{\|k\|} \|x - y\|_{\mathbb{X}}$ .  $\square$

In particular, this means that  $\tau_{k,l} : \mathbb{X} \rightarrow \mathbb{X}$  is continuous in the topology of pointwise convergence: if  $x^n \rightarrow x^\infty$  pointwise, then  $\tau_{k,l}x^n \rightarrow \tau_{k,l}x^\infty$  pointwise. Of course, this is also clear without Lemma 5.3.

The main result of this section is the following theorem, which says that under the condition that the local potentials  $S_j$  are twice continuously differentiable with uniformly bounded second derivatives, then  $-\nabla W$  indeed defines a flow on  $\mathbb{X}$ . Moreover, this flow has the regularity properties one expects it to have.

**Theorem 5.4.** Assume the local potentials  $S_j$  satisfy conditions A, B and E, that is they are twice continuously differentiable with uniformly bounded second derivatives, they depend on finitely many variables and are shift-invariant. Then the vector field  $-\nabla W : \mathbb{X} \rightarrow \mathbb{X}$  is globally Lipschitz continuous, i.e. there is a constant  $L > 0$ , depending only on the constant  $C$  of condition E and the interaction range  $r$  of condition A, such that for all  $x, y \in \mathbb{X}$ ,

$$\|\nabla W(x) - \nabla W(y)\|_{\mathbb{X}} \leq L \|x - y\|_{\mathbb{X}}.$$

Hence, the initial value problem  $\frac{dx}{dt} = -\nabla W(x)$ ,  $x(0) = x_0$  on  $\mathbb{X}$  has global-in-time existence and uniqueness of solutions and defines a complete flow  $t \mapsto \Psi_t$  on  $\mathbb{X}$ . This flow is Lipschitz continuous, i.e. there are constants  $L_t > 0$ , depending only on  $L$ , such that for all  $t \in \mathbb{R}$  and  $x, y \in \mathbb{X}$ ,

$$\|\Psi_t x - \Psi_t y\|_{\mathbb{X}} \leq L_t \|x - y\|_{\mathbb{X}}.$$

Moreover, this flow depends Lipschitz continuously on  $-\nabla W$ . This means that there are constants  $\bar{L}_t > 0$ , depending only on  $L$ , such that for all  $t \in \mathbb{R}$  and  $x, y \in \mathbb{X}$  and for all  $-\nabla W$  and  $-\nabla \tilde{W}$  with Lipschitz constants  $\leq L$  and respective complete flows  $\Psi_t$  and  $\tilde{\Psi}_t$ ,

$$\sup_{x \in \mathbb{X}} \|\Psi_t x - \tilde{\Psi}_t x\|_{\mathbb{X}} \leq \bar{L}_t \sup_{x \in \mathbb{X}} \|\nabla W(x) - \nabla \tilde{W}(x)\|_{\mathbb{X}}.$$

**Proof.** Using the uniform bound that  $|\partial_{i,k} S_j| \leq C$ , see condition E, we will prove that  $-\nabla W$  maps  $\mathbb{X}$  to  $\mathbb{X}$  and is globally Lipschitz continuous. The usual ODE theory then provides the existence of a complete flow  $t \mapsto \Psi_t$  on  $\mathbb{X}$ .

Thus, let  $x, y \in \mathbb{X}$ . Then first of all

$$\begin{aligned} |-\nabla W(x)_i + \nabla W(y)_i| &\leq \sum_{\|j-i\| \leq r} |\partial_i S_j(y) - \partial_i S_j(x)| = \sum_{\|j-i\| \leq r} \left| \int_0^1 \frac{d}{d\tau} (\partial_i S_j(\tau y + (1-\tau)x)) d\tau \right| \\ &\leq \sum_{\|j-i\| \leq r} \left| \int_0^1 \sum_{\|k-j\| \leq r} \partial_{i,k} S_j(\tau y + (1-\tau)x) d\tau \right| \cdot |y_k - x_k| \\ &\leq C \sum_{\|k-j\| \leq r} \sum_{\|j-i\| \leq r} |x_k - y_k|. \end{aligned}$$

But this implies that

$$\begin{aligned} \|-\nabla W(x) + \nabla W(y)\|_{\mathbb{X}} &\leq C \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|} \sum_{\|k-j\| \leq r} \sum_{\|j-i\| \leq r} |x_k - y_k| \\ &= C \sum_{\|m\| \leq r} \sum_{\|n\| \leq r} \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|} |x_{i+m+n} - y_{i+m+n}| \\ &= C \sum_{\|m\| \leq r} \sum_{\|n\| \leq r} \|\tau_{m+n,0} x - \tau_{m+n,0} y\|_{\mathbb{X}} \end{aligned}$$

By Lemma 5.3 and the fact that in the sum above  $\|m+n\| \leq 2r$ , we know that  $\|\tau_{m+n,0} x - \tau_{m+n,0} y\|_{\mathbb{X}} \leq 2^{2r} \|x - y\|_{\mathbb{X}}$ . Hence, noting that  $|\{i \in \mathbb{Z}^d \mid \|i\| \leq r\}| \leq (2r+1)^d$ , we obtain that

$$\|-\nabla W(x) + \nabla W(y)\|_{\mathbb{X}} \leq L \|x - y\|_{\mathbb{X}},$$

where  $L := 2^{2r} C (2r+1)^{2d}$ . On the one hand, this shows that  $-\nabla W$  is globally Lipschitz continuous. On the other hand, choosing  $y = 0$ , we see that  $\|-\nabla W(x) + \nabla W(0)\|_{\mathbb{X}} \leq L \|x\|_{\mathbb{X}}$ , or  $\|-\nabla W(x)\|_{\mathbb{X}} \leq L \|x\|_{\mathbb{X}} + \|\nabla W(0)\|_{\mathbb{X}}$ , that is  $-\nabla W$  maps  $\mathbb{X}$  into  $\mathbb{X}$ .

This implies the existence and uniqueness of solutions of the initial value problem  $\frac{dx}{dt} = -\nabla W(x)$ ,  $x(0) = x_0$  in  $\mathbb{X}$ , that is the existence of flow maps  $\Psi_t : \mathbb{X} \rightarrow \mathbb{X}$  for all  $t \in \mathbb{R}$ . The Lipschitz continuity of  $\Psi_t$  follows, as usual, from an application of Gronwall's inequality: first one shows that  $\|\Psi_t x - \Psi_t y\|_{\mathbb{X}} \leq \|x - y\|_{\mathbb{X}} + L \int_0^{|t|} \|\Psi_\tau x - \Psi_\tau y\|_{\mathbb{X}} d\tau$ . This then implies that  $\|\Psi_t x - \Psi_t y\|_{\mathbb{X}} \leq L_t \|x - y\|_{\mathbb{X}}$ , with  $L_t = e^{L|t|}$ .

For the last part of the theorem, let  $\nabla W$  and  $\nabla \tilde{W}$  be two vector fields with Lipschitz constants  $\leq L$  and complete flows  $\Psi_t$  and  $\tilde{\Psi}_t$  respectively. Call  $x(t) = \Psi_t x$  and  $\tilde{x}(t) = \tilde{\Psi}_t x$ . We then have

$$\begin{aligned}
\|x(t) - \tilde{x}(t)\|_{\mathbb{X}} &\leq \int_0^{|t|} \|\nabla \tilde{W}(\tilde{x}(\tau)) - \nabla W(x(\tau))\|_{\mathbb{X}} d\tau \\
&\leq \int_0^{|t|} \|\nabla \tilde{W}(\tilde{x}(\tau)) - \nabla W(\tilde{x}(\tau))\|_{\mathbb{X}} d\tau + \int_0^{|t|} \|\nabla W(\tilde{x}(\tau)) - \nabla W(x(\tau))\|_{\mathbb{X}} d\tau \\
&\leq |t| \sup_{x \in \mathbb{X}} \|\nabla W(x) - \nabla \tilde{W}(x)\|_{\mathbb{X}} + L \int_0^{|t|} \|\tilde{x}(\tau) - x(\tau)\|_{\mathbb{X}} d\tau.
\end{aligned}$$

Thus, by Gronwall's inequality,  $\|\psi_t x - \tilde{\psi}_t x\|_{\mathbb{X}} \leq \bar{L}_t \sup_{x \in \mathbb{X}} \|\nabla W(x) - \nabla \tilde{W}(x)\|_{\mathbb{X}}$  with  $\bar{L}_t = |t|e^{L|t|}$ .  $\square$

**Remark 5.5.** It is not true in general that  $-\nabla W : \mathbb{X} \rightarrow \mathbb{X}$  is a  $C^1$  map. Hence, contrary to a claim made in [9], the  $\psi_t$  in general cannot be assumed  $C^1$  either.

By Proposition 5.2, the first part of Theorem 5.4 implies that  $\nabla W : \mathbb{X} \rightarrow \mathbb{X}$  is continuous with respect to pointwise convergence: if  $\lim_{n \rightarrow \infty} x^n = x^\infty$  pointwise, then  $\lim_{n \rightarrow \infty} \nabla W(x^n) = \nabla W(x^\infty)$  pointwise.

Similarly, the second part of Theorem 5.4 implies that for every  $t \in \mathbb{R}$  the flow map  $\psi_t : \mathbb{X} \rightarrow \mathbb{X}$  is continuous with respect to pointwise convergence.

Part three of Theorem 5.4 implies that if  $\nabla W^n, \nabla W^\infty : \mathbb{X} \rightarrow \mathbb{X}$  is a sequence of formal gradient vector fields with a uniform Lipschitz constant and corresponding flow maps  $\psi_t^n, \psi_t^\infty : \mathbb{X} \rightarrow \mathbb{X}$  and such that  $\nabla W^n \rightarrow \nabla W^\infty$  uniformly on  $\mathbb{X}$ , then for all  $t$  also  $\psi_t^n \rightarrow \psi_t^\infty$  uniformly on  $\mathbb{X}$ .

In the remainder of this section, we will formulate a concept of convergence for a sequence of finite range potentials  $S_j^n$  that guarantees that their corresponding gradient vector fields and flow maps converge uniformly. It turns out that it is enough to require the convergence of the gradients of the  $S_j^n$ . We will first of all need to define what it means for collections of gradients of finite range potentials to be “close”. Remembering the definition in Section 2 of the partial derivatives  $\partial_{j_1, \dots, j_k} S$  of a  $k$  times continuously differentiable function  $S : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  of finitely many variables, we now define:

**Definition 5.6.** Let  $S : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  be an  $m+1 \geq 0$  times continuously differentiable function of finitely many variables, that is  $S(x) = s(x|_B)$  for a certain finite subset  $B \subset \mathbb{Z}^d$  and an  $m+1$  times continuously differentiable function  $s : \mathbb{R}^B \rightarrow \mathbb{R}$ . Then we define the uniform  $C^m(\mathbb{R}^{\mathbb{Z}^d})$  norm  $\|\nabla S\|_{C^m(\mathbb{R}^{\mathbb{Z}^d})} \in [0, \infty)$  of the gradient of  $S$  as the finite sum of suprema

$$\|\nabla S\|_{C^m(\mathbb{R}^{\mathbb{Z}^d})} := \sum_{1 \leq k \leq m+1} \sum_{j_1, \dots, j_k \in \mathbb{Z}^d} \sup_{x \in \mathbb{R}^{\mathbb{Z}^d}} |\partial_{j_1, \dots, j_k} S(x)|.$$

We note that if  $S_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  is a collection of  $m+1$  times continuously differentiable, shift invariant finite range potentials, that is if  $S_j(x) = s_j(x|_{B_j^r})$  for some  $m+1$  times continuously differentiable function  $s_j : \mathbb{R}^{B_j^r} \rightarrow \mathbb{R}$  and  $S_j(\tau_{k,l} x) = S_{j+k}(x)$  for all  $j, k$  and  $l$ , then  $\|\nabla S_i\|_{C^m(\mathbb{R}^{\mathbb{Z}^d})} = \|\nabla S_j\|_{C^m(\mathbb{R}^{\mathbb{Z}^d})}$  for all  $i, j \in \mathbb{Z}^d$ . With this in mind, we first of all prove:

**Proposition 5.7.** Let  $S_j, \tilde{S}_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  be two  $m+1 \geq 1$  times continuously differentiable and shift-invariant collections of finite range local potentials, say  $S_j(x) = s_j(x|_{B_j^r})$  and  $\tilde{S}_j(x) = \tilde{s}_j(x|_{B_j^r})$  and denote their cor-

responding gradient vector fields by  $\nabla W$  and  $\nabla \tilde{W}$ . Then, there is a constant  $\tilde{L} > 0$ , depending only on the dimension  $d$ , such that

$$\sup_{x \in \mathbb{X}} \|\nabla W(x) - \nabla \tilde{W}(x)\|_{\mathbb{X}} \leq \tilde{L} \|\nabla S_0 - \nabla \tilde{S}_0\|_{C^0(\mathbb{R}^d)}.$$

**Proof.** We have that

$$\|\nabla W(x) - \nabla \tilde{W}(x)\|_{\mathbb{X}} \leq \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|} \sum_{\|j-i\| \leq r} |\partial_i S_j(x) - \partial_i \tilde{S}_j(x)|.$$

By shift invariance,  $\partial_i S_j(x) = \partial_{i-j} S_0(\tau_{j,0}x)$  and similarly for  $\tilde{S}_j$ , so that  $\sup_x |\partial_i S_j(x) - \partial_i \tilde{S}_j(x)| = \sup_x |\partial_{i-j} S_0(x) - \partial_{i-j} \tilde{S}_0(x)|$ , and consequently

$$\begin{aligned} \sup_{x \in \mathbb{X}} \|\nabla W(x) - \nabla \tilde{W}(x)\|_{\mathbb{X}} &\leq \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|} \sum_{\|j-i\| \leq r} \sup_{x \in \mathbb{X}} |\partial_{i-j} S_0(x) - \partial_{i-j} \tilde{S}_0(x)| \\ &= \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|} \sum_{\|j\| \leq r} \sup_{x \in \mathbb{X}} |\partial_j S_0(x) - \partial_j \tilde{S}_0(x)| \\ &\leq \tilde{L} \|\nabla S_0 - \nabla \tilde{S}_0\|_{C^0(\mathbb{R}^d)}, \end{aligned}$$

with  $\tilde{L} = \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|}$ .  $\square$

We are now ready to define what it means for a sequence of local potentials to converge:

**Definition 5.8.** Let  $S_j^n, S_j^\infty : \mathbb{R}^d \rightarrow \mathbb{R}$  be a sequence of collections of  $m+1 \geq 1$  times continuously differentiable, shift-invariant functions of finite range  $r$ . Then we say that the  $\nabla S_j^n$  converge to the  $\nabla S_j^\infty$  uniformly in  $C^m(\mathbb{R}^d)$  as  $n \rightarrow \infty$  if

$$\lim_{n \rightarrow \infty} \|\nabla S_0^n - \nabla S_0^\infty\|_{C^m(\mathbb{R}^d)} = 0.$$

With this definition, we can then prove the following corollary of Theorem 5.4. It trivially follows from our definitions, Theorem 5.4 and Proposition 5.7.

**Corollary 5.9.** Let  $S_j^n, S_j^\infty : \mathbb{R}^d \rightarrow \mathbb{R}$  be a sequence of continuously differentiable local potentials of finite range  $r$ , with corresponding gradient vector fields  $\nabla W^n$  and  $\nabla W^\infty$ , and assume that  $\nabla S_j^n \rightarrow \nabla S_j^\infty$  uniformly in  $C^0(\mathbb{R}^d)$ . Then  $\nabla W^n \rightarrow \nabla W^\infty$  uniformly, i.e.

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{X}} \|\nabla W^n(x) - \nabla W^\infty(x)\|_{\mathbb{X}} = 0.$$

Moreover, in the case that the  $S_j^n$  and  $S_j^\infty$  are twice continuously differentiable with uniformly bounded second derivatives, so that  $-\nabla W^n$  and  $-\nabla W^\infty$  have well-defined flow maps  $\Psi_t^n$  and  $\Psi_t^\infty$ , then it also holds for every  $t \in \mathbb{R}$  that  $\Psi_t^n \rightarrow \Psi_t^\infty$  uniformly, i.e.

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{X}} \|\Psi_t^n x - \Psi_t^\infty x\|_{\mathbb{X}} = 0.$$



## 6. Properties of the gradient flow

In this section, we collect some qualitative properties of the formal negative gradient flow that was introduced in the previous section.

First of all, not surprisingly, it is equivariant with respect to shifts:

**Proposition 6.1.** *Let  $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}$  and  $t \in \mathbb{R}$ . Then  $\Psi_t \circ \tau_{k,l} = \tau_{k,l} \circ \Psi_t$ .*

**Proof.** By the shift-invariance of the local potentials  $S_j$  of condition B above, we have that  $S_{j-k}(\tau_{k,l}x) = S_j(x)$  for all  $k, l$  and  $j$ . Differentiating this identity with respect to  $x_{i+k}$ , we find that  $\partial_i S_{j-k}(\tau_{k,l}x) = \partial_{i+k} S_j(x)$ . Assume now that  $\frac{dx_i}{dt} = -(\nabla W(x))_i$  for all  $i$ . Then,

$$\begin{aligned} \frac{d}{dt}(\tau_{k,l}x)_i &= \frac{dx_{i+k}}{dt} = -(\nabla W(x))_{i+k} = - \sum_{\|j-(i+k)\| \leq r} \partial_{i+k} S_j(x) \\ &= - \sum_{\|(j-k)-i\| \leq r} \partial_i S_{j-k}(\tau_{k,l}x) = -(\nabla W(\tau_{k,l}x))_i. \end{aligned}$$

In other words, when  $t \mapsto x(t)$  is a solution of the negative gradient flow, then so is  $t \mapsto \tau_{k,l}x(t)$ .  $\square$

Proposition 6.1 implies in particular that the spaces  $\mathbb{X}_{p,q}$  of periodic configurations are invariant under the gradient flow.

The following well-known property of the negative gradient flow is the analogue of the comparison principle for parabolic PDEs, cf. [7] or [11]. It is a direct consequence of the monotonicity condition D.

**Theorem 6.2** (Strict monotonicity of the parabolic flow). *Let  $x, y \in \mathbb{X}$  such that  $x < y$ . Denote by  $\Psi_t$  the time- $t$  flow of  $\dot{x} = -\nabla W(x)$ . Then for every  $t > 0$ ,  $\Psi_t x \ll \Psi_t y$ .*

**Proof.** Denote  $x(t) = \Psi_t x$  and  $y(t) = \Psi_t y$  and define  $u(t) := y(t) - x(t)$ . Note that  $u(0) > 0$  and that  $u$  satisfies the following linear ODE:

$$\begin{aligned} \dot{u}_i(t) &= -\partial_i W(y(t)) + \partial_i W(x(t)) = \int_0^1 \frac{d}{d\tau} \left( \sum_{\|i-j\| \leq r} \partial_i S_j(\tau x(t) + (1-\tau)y(t)) \right) d\tau \\ &= \sum_{\|i-j\| \leq r, \|j-k\| \leq r} \left( \int_0^1 -\partial_{i,k} S_j(\tau x(t) + (1-\tau)y(t)) d\tau \right) u_k(t) =: (H(t)u(t))_i. \end{aligned}$$

Here, for every  $t$ , the operator  $H(t)$  is Lipschitz from  $\mathbb{X}$  to  $\mathbb{X}$ , by a proof similar to that of Theorem 5.4. Recall that  $\partial_{i,k} S_j \leq 0$  when  $i \neq k$ , whereas  $\partial_{i,i} S_j < C$ . This implies that there is a constant  $M > 0$  such that the operators  $\tilde{H}(t) := H(t) + M \text{Id} : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$  are positive:  $u \geq 0$  implies  $\tilde{H}(t)u \geq 0$ .

Note moreover that both the  $H(t)$  and the  $\tilde{H}(t)$  are uniformly bounded operators, whence the ODEs  $\dot{u} = H(t)u$  and  $\dot{v} = \tilde{H}(t)v$  define well-posed initial value problems. More importantly,  $u(t)$  solves  $\dot{u} = H(t)u$  if and only if  $v(t) := e^{Mt}u(t)$  solves  $\dot{v} = \tilde{H}(t)v$ . We will now prove that for every  $t > 0$  and every  $i$ ,  $v_i(t) > 0$ . Then, obviously,  $u_i(t) > 0$  as well, which then proves the theorem.

To prove the claim on  $v(t)$ , we solve the initial value problem for  $\dot{v} = \tilde{H}(t)v$  by Picard iteration, that is we write

$$v(t) = \left( \sum_{n=0}^{\infty} \tilde{H}^{(n)}(t) \right) v(0), \quad (6.12)$$

where the  $\tilde{H}^{(n)}(t)$  are defined inductively by

$$\tilde{H}^{(0)}(t) = \text{Id} \quad \text{and} \quad \tilde{H}^{(n)}(t) := \int_0^t \tilde{H}(\tilde{t}) \circ \tilde{H}^{(n-1)}(\tilde{t}) d\tilde{t} \quad \text{for } n \geq 1.$$

Observe that the positivity of  $\tilde{H}(t)$  implies that the  $\tilde{H}^{(n)}(t)$  are positive as well. Because  $v(0) = u(0) > 0$ , we can therefore estimate, for any  $i, k \in \mathbb{Z}^d$  with  $\|i - k\| = 1$ ,

$$\begin{aligned} v_i(t) &= \left( \sum_{n=0}^{\infty} \tilde{H}^{(n)}(t) v(0) \right)_i \geq (\tilde{H}(t) v(0))_i \\ &\geq \left( \int_0^t \int_0^1 -\partial_{i,k} S_i(\tau x(\tilde{t}) + (1-\tau)y(\tilde{t})) d\tau d\tilde{t} \right) v_k(0). \end{aligned} \quad (6.13)$$

Now choose a  $k \in \mathbb{Z}^d$  such that  $v_k(0) > 0$  and recall that  $\partial_{i,k} S_i < 0$ . Then from (6.13) it follows that if  $\|i - k\| = 1$ , then for all  $t > 0$ ,  $v_i(t) > 0$ .

To generalize to the case that  $\|i - k\| \neq 1$ , let us choose a sequence of lattice points  $k = i_0, \dots, i_N = i$  such that  $\|i_n - i_{n-1}\| = 1$  and  $N = \|i - k\|$ . Then, by induction,  $v(\frac{nt}{N})_{i_n} > 0$  for all  $n$ . Thus, if  $v_k(0) > 0$  and  $t > 0$ , then  $v_i(t) > 0$ .  $\square$

Theorem 6.2 immediately gives us the following important corollary.

**Corollary 6.3.** *Let  $\omega \in \mathbb{R}^d$ . Then  $\mathcal{B}_\omega$  is positively invariant under the negative gradient flow:  $\Psi_t(\mathcal{B}_\omega) \subset \mathcal{B}_\omega$ , for every  $t > 0$ .*

This just follows because the strict monotonicity of the parabolic flow implies that  $\Psi_t$  preserves the inequalities that define  $\mathcal{B}_\omega$ .

The following is a quantitative version of Theorem 6.2. It will be crucial in the remainder of this paper and we have not found it elsewhere in the literature.

**Theorem 6.4** (Parabolic Harnack inequality). *Let  $t > 0$ ,  $K \subset \mathbb{R}^d$  a compact set and  $x, y \in \mathcal{B}_K := \bigcup_{\omega \in K} \mathcal{B}_\omega$  such that  $x < y$ . Then there exists a constant  $L > 0$ , depending only on  $K$ ,  $\|i - k\|$  and  $t$ , such that for all  $i, k \in \mathbb{Z}^d$ ,*

$$(\Psi_t y)_i - (\Psi_t x)_i \geq L(y_k - x_k).$$

**Proof.** The proof is a quantitative variant of the proof of Theorem 6.2. We start by recalling that by Corollary 3.7, there is a constant  $\lambda > 0$ , depending only on  $K$ , such that  $\partial_{i,k} S_i(z) \leq -\lambda < 0$  for all  $\|i - k\| = 1$  and  $z \in \bigcup_{\omega \in K} \mathcal{B}_\omega$ . Then (6.13) shows that if  $\|i - k\| = 1$ , then  $v_i(t) \geq \tilde{L}_1 v_k(0)$ , with  $\tilde{L}_1 = t\lambda$ .

To generalize to the case that  $\|k - i\| \neq 1$ , we again choose a sequence of lattice points  $k = i_0, \dots, i_N = i$  such that  $\|i_n - i_{n-1}\| = 1$  and  $N = \|i - k\|$ . Then there is a constant  $\tilde{L}'_N = \frac{t\lambda}{N}$  depending only on  $K$ ,  $t$  and  $\|i - k\|$  such that  $v(\frac{nt}{N})_{i_n} \geq \tilde{L}'_N v(\frac{(n-1)t}{N})_{i_{n-1}}$  for all  $n$ . Thus,  $v_i(t) \geq \tilde{L}_N v_k(0)$  with  $\tilde{L}_N = (\tilde{L}'_N)^N$ .

This proves that  $u_i(t) \geq Lu_k(0)$  with  $L = e^{-Mt} \tilde{L}_N = e^{-Mt} (\lambda t / \|i - k\|)^{\|i - k\|}$ .  $\square$

Note that for Birkhoff configurations, both the strict monotonicity, Theorem 6.2, and the elliptic Harnack inequality, Theorem 4.12, follow directly from this parabolic Harnack inequality.

We moreover remark that under the uniform twist condition that  $\partial_{i,k} S_i(z) \leq -\lambda < 0$  for all  $z \in \mathbb{X}$  and  $\|i - k\| = 1$ , the above parabolic Harnack inequality holds for all  $x, y$  in  $\mathbb{X}$  with  $x < y$ , i.e. it then

holds irrespective of the Birkhoff property of  $x$  and  $y$ . This uniform twist condition for instance holds for the Frenkel–Kontorova problem, see formula (1.5).

To finish this section, let us for completeness include the following alternative existence proof for globally stationary Birkhoff solutions of arbitrary rotation vector. It was provided by Golé in [7] in dimension  $d = 1$ . The below is a more or less trivial generalization to higher dimensions, see also [11]. As opposed to the results presented in Section 4.2, it also holds without hypothesis C that requires that the  $S_j$  are coercive. The proof presented here is slightly shorter and more direct than the proof in [11].

**Theorem 6.5.** *Also without the coercivity condition C, it holds that for every  $\omega \in \mathbb{R}^d$ , there exists an  $x \in \mathcal{B}_\omega$  with  $\nabla W(x) = 0$ .*

**Proof.** Recall that the conditions A, D and E alone guarantee that the compact set  $\mathcal{B}_\omega$  is forward invariant under the negative gradient flow. Condition B will be used below.

Now, for  $B \subset \mathbb{Z}^d$  a finite subset, recall the definition of the finite action  $W_B(x) := \sum_{j \in B} S_j(x)$ . Then, for  $i \in \hat{B}^{(r)}$ , it holds that  $\partial_i W_B(x) = \partial_i W(x)$ , whereas if  $\|i - B\| := \min_{j \in B} \|j - i\| > r$ , it is true that  $\partial_i W_B(x) = 0$ . Thus, the time-derivative of  $W_B$  along solutions of  $\frac{dx}{dt} = -\nabla W(x)$  equals

$$\frac{d}{dt} W_B(x) = - \sum_{i \in \mathbb{Z}^d} \partial_i W_B(x) \partial_i W(x) = - \sum_{i \in \hat{B}^{(r)}} (\partial_i W(x))^2 - \sum_{i \notin \hat{B}^{(r)}, \|i - B\| \leq r} \partial_i W_B(x) \partial_i W(x).$$

We call  $A_B(x) := \sum_{i \in \hat{B}^{(r)}} (\partial_i W(x))^2$ . It is the square length of the gradient of the map  $y \mapsto W_B(x + y)$  from  $\mathbb{R}^{\hat{B}^{(r)}}$  to  $\mathbb{R}$  evaluated at  $y = 0$ . With this definition, one checks that if  $B_1 \subset B_2$ , then  $A_{B_1}(x) \leq A_{B_2}(x)$ . Moreover, if  $\hat{B}_1^{(r)}$  and  $\hat{B}_2^{(r)}$  are disjoint, then  $A_{B_1 \cup B_2}(x) = A_{B_1}(x) + A_{B_2}(x)$ .

The second sum in the expression for  $\frac{d}{dt} W_B$  consists of “boundary terms”. We will call it  $a_B(x) := \sum_{i \notin \hat{B}^{(r)}} \partial_i W_B(x) \partial_i W(x)$ . Because  $\mathcal{B}_\omega$  is compact and  $\partial_i S_j = \partial_{i-j} S_0 \circ \tau_{j,0}$  for all  $j \in \mathbb{Z}^d$ , there is a constant  $c > 0$  with the property that  $|\partial_i S_j| \leq c$  for all  $i, j \in \mathbb{Z}^d$  and uniformly on  $\mathcal{B}_\omega$ . This in turn implies the estimate  $|a_B(x)| \leq c^2 (2r + 1)^d |\partial B|$ .

Assume now that there is no globally stationary point in  $\mathcal{B}_\omega$ . Then for every  $x \in \mathcal{B}_\omega$  there is a finite subset  $B_x \subset \mathbb{Z}^d$  such that  $A_{B_x}(x) = 2\varepsilon_x > 0$ . Moreover, because  $\nabla W_{B_x}$  is continuous on  $\mathbb{R}^{\mathbb{Z}^d}$ , it holds that  $x$  has an open neighborhood  $U_x \subset \mathbb{R}^{\mathbb{Z}^d}$  on which  $A_{B_x} > \varepsilon_x$ . By compactness we can find a finite collection  $x_1, \dots, x_m \in \mathcal{B}_\omega$  such that  $\mathcal{B}_\omega \subset \bigcup_{i=1}^m U_{x_i}$ . Define  $B := \bigcup_{i=1}^m B_{x_i}$  and  $\varepsilon := \min_i \{\varepsilon_{x_i}\} > 0$ . Then every  $x \in \mathcal{B}_\omega$  is in some  $U_{x_i}$  and thus,  $A_B(x) \geq A_{B_{x_i}}(x) > \varepsilon_{x_i} \geq \varepsilon > 0$ , that is  $A_B > \varepsilon > 0$  uniformly on  $\mathcal{B}_\omega$ . Moreover, translation invariance implies that for any  $k \in \mathbb{Z}^d$ , also  $A_{B+k} > \varepsilon$  uniformly on  $\mathcal{B}_\omega$ .

For  $n \in \mathbb{N}$ , define the ball  $B(n) := \{j \in \mathbb{Z}^d \mid \|j\| \leq n\} \subset \mathbb{Z}^d$  and let  $N \in \mathbb{N}$  be such that the  $B$  above is contained in  $B(N)$ . Then  $A_{B(N)}(x) \geq A_B(x) > \varepsilon > 0$ .

Let  $m \geq 2$  be an integer. By translation invariance and the fact that  $B(mN)$  contains at least  $m^d$  translates of  $B(N)$  with disjoint  $r$ -interiors, it holds that  $A_{B(mN)} \geq m^d A_{B(N)} > m^d \varepsilon$ . On the other hand,  $|a_{mN}(x)| \leq c^2 (2r + 1)^d |\partial B(mN)| = D_N^r m^{d-1}$  for some  $D_N^r > 0$ . Thus,  $\frac{d}{dt} W_{B(mN)}(x) \leq -m^d \varepsilon + D_N^r m^{d-1}$  and hence by choosing  $m$  large enough, we can arrange that  $\frac{d}{dt} W_{mN}(x) \leq -1$  uniformly on  $\mathcal{B}_\omega$ .

Since  $\mathcal{B}_\omega$  is forward invariant under the negative gradient flow, this implies that  $W_{mN}$  is not bounded from below on  $\mathcal{B}_\omega$ . This contradicts the fact that  $\mathcal{B}_\omega$  is compact and  $W_{mN}$  is continuous. This proves that there must be a globally stationary point in  $\mathcal{B}_\omega$ .  $\square$

## 7. Ghost circles

In dimension  $d = 1$ , the concept of a ghost circle was introduced by Golé. We generalize this definition here to general dimensions. Note the similarity with Definition 4.13 of an Aubry–Mather set.

**Definition 7.1** (*Ghost circle*). A ghost circle  $\Gamma \subset \mathbb{R}^{\mathbb{Z}^d}$  is a collection of configurations with the following properties:

- $\Gamma$  is nonempty, closed and connected.
- $\Gamma$  is strictly ordered, i.e. for every  $x, y \in \Gamma$ ,  $x \ll y$ ,  $x = y$  or  $x \gg y$ .
- $\Gamma$  is invariant under shifts: if  $x \in \Gamma$ , then for every  $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}$ , also  $\tau_{k,l}x \in \Gamma$ .
- $\Gamma$  is invariant under the positive and negative gradient flow: for all  $t \in \mathbb{R}$ ,  $\psi_t(\Gamma) = \Gamma$ .

An example of a ghost circle are the connected Aubry–Mather sets of Theorem 4.18.

The strict ordering and the shift-invariance of a ghost circle  $\Gamma$  imply that any configuration  $x \in \Gamma$  is Birkhoff and hence has a rotation vector  $\omega = \omega(x)$ . The ordering of  $\Gamma$  moreover implies that this rotation vector is independent of the choice of  $x \in \Gamma$ , that is  $\omega = \omega(\Gamma)$  and thus,  $\Gamma \subset \mathcal{B}_\omega$ .

Let  $j \in \mathbb{Z}^d$ . Recall the definition of the projection to the  $j$ -th factor

$$\pi_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}, \quad \pi_j(x) = x_j.$$

Each  $\pi_j$  is continuous with respect to pointwise convergence. In fact, we can show that  $\pi_j|_\Gamma : \Gamma \rightarrow \mathbb{R}$  is a homeomorphism:

**Proposition 7.2.** *Let  $\Gamma$  be a ghost circle. Then, for every  $j \in \mathbb{Z}^d$ , the projection  $\pi_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  induces a homeomorphism  $\pi_j|_\Gamma : \Gamma \rightarrow \mathbb{R}$ .*

**Proof.** Let  $\Gamma \subset \mathbb{R}^{\mathbb{Z}^d}$  be a ghost circle. Clearly,  $\pi_j|_\Gamma$  is continuous.

The strict ordering of  $\Gamma$  implies that  $\pi_j|_\Gamma$  is injective. Moreover, shift-invariance of  $\Gamma$  implies that if  $x \in \Gamma$ , then so is  $\tau_{0,l}x = x + l$  for every  $l \in \mathbb{Z}$ , whence the range of  $\pi_j|_\Gamma$  is unbounded. Since  $\Gamma$  is connected and  $\pi_j|_\Gamma$  is continuous, its range is both unbounded and connected, that is  $\pi_j|_\Gamma : \Gamma \rightarrow \mathbb{R}$  is surjective.

To prove that  $(\pi_j|_\Gamma)^{-1} : \mathbb{R} \rightarrow \Gamma$  is continuous, it suffices to realize that  $\pi_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  is an open map, i.e. that it sends open sets to open sets. This holds because the topology of pointwise convergence is generated by open sets  $U \subset \mathbb{R}^{\mathbb{Z}^d}$  for which  $\pi_k(U) = V$  with  $V \subset \mathbb{R}$  an open subset, while  $\pi_l(U) = \mathbb{R}$  for all  $l \neq k$ . For such  $U$ , it is clear that  $\pi_j(U)$  is open.  $\square$

Lemma 7.2 thus says that a ghost circle  $\Gamma$  is homeomorphic to  $\mathbb{R}$ . It should be remarked though that, because  $\Gamma$  is invariant under the vertical shift  $\tau_{0,1}$ , and the gradient flow  $\psi_t$  is equivariant with respect to  $\tau_{0,1}$ , it makes sense to identify every element  $x \in \Gamma$  with  $\tau_{0,1}x = x + 1 \in \Gamma$ . The quotient  $\Gamma/\mathbb{Z} \cong \mathbb{R}/\mathbb{Z}$  is a genuine topological circle. This identification is sometimes understood in this paper.

The name *ghost circle* refers to the fact that  $\Gamma/\mathbb{Z}$  may not consist of “physically relevant” configurations, i.e. globally stationary solutions. But, being a compact one-dimensional object consisting of orbits of a formal gradient flow, it has a good chance of containing such solutions. In fact, the following proposition serves as a first motivation to study ghost circles.

**Proposition 7.3.** *Every ghost circle  $\Gamma \subset \mathbb{R}^{\mathbb{Z}^d}$  contains a globally stationary solution.*

Since  $\Gamma$  is a closed, flow-invariant subset of some  $\mathcal{B}_\omega$ , the proof of this proposition is identical to that of Theorem 6.5. Moreover, we remark that when  $\Gamma$  contains at least one global minimizer, say  $x$ , then it automatically contains the entire Aubry–Mather set  $\mathcal{M}(x)$ .

In the following two sections we first of all show that under generic conditions, ghost circles of rational rotation vectors exist and then we will prove a compactness result for ghost circles which will allow us to take limits and obtain ghost circles of irrational rotation vectors.

## 8. Morse approximations and periodic ghost circles

In this section, we will prove two technical results. The first is that the local potentials  $S_j$  can be perturbed, in a way that will be made precise, so that the periodic action  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  becomes a Morse function.

The second result of this section says that whenever  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is a Morse function, then there exists a ghost circle  $\Gamma \subset \mathbb{X}_{p,q}$ .

Together with the results of Section 9, this will imply that any collection of local potentials admits a ghost circle of arbitrary rotation vector.

### 8.1. Existence of Morse approximations

Let  $\omega \in \mathbb{Q}^d$  be a rational rotation vector and let  $(p_1, q_1), \dots, (p_d, q_d)$  be a set of principal periods for  $\omega$ . Recall that in Section 4 we defined the periodic action function  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  by  $W_{p,q}(x) = \sum_{j \in B_p} S_j(x)$ .

One says that  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is a *Morse function* if at its critical points its Hessian is nondegenerate. In other words, if  $\nabla W_{p,q}(x) = 0$  implies that  $D^2 W_{p,q}(x)$  is invertible, where  $D^2 W_{p,q}(x)$  is the symmetric matrix of second derivatives of  $W_{p,q}$  evaluated at  $x$ . By the implicit function theorem, every critical point  $x$  of a Morse function is isolated. Moreover, each of these critical points can be assigned an index  $i(x)$  which equals the dimension of the unstable manifold of  $x$ , considered as an equilibrium point for the negative gradient flow  $\frac{dx}{dt} = -\nabla W_{p,q}(x)$ .

We remark here that for arbitrary local potentials  $S_j$ , the periodic action  $W_{p,q}$  is not automatically a Morse function. A simple example of a non-Morse action function arises in the Frenkel–Kontorova model without local potential, for which

$$W_{p,q}(x) = \sum_{j \in B_p} \frac{1}{8d} \sum_{\|i-j\|=1} (x_i - x_j)^2.$$

This action function satisfies  $W_{p,q}(x+t) = W_{p,q}(x)$  for all  $t \in \mathbb{R}$ , so that its second derivative is everywhere degenerate. In fact, it has a one-parameter family of stationary points, and thus none of those is isolated. Nevertheless, in this subsection we will prove the following theorem:

**Theorem 8.1.** *Let  $S_j : \mathbb{R}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  be local potentials that satisfy conditions A–E. Let  $\omega \in \mathbb{Q}^d$  and let  $(p_1, q_1), \dots, (p_d, q_d)$  be principal periods  $\omega$ , that is  $\overline{\mathbb{X}}_\omega = \mathbb{X}_{p,q}$ . Then there exists a sequence of local potentials  $S_j^n$  with the following properties:*

1. The  $S_j^n$  satisfy conditions A–E.
2. The range of interaction of the  $S_j^n$  is uniformly bounded in  $n$ .
3. For every  $n$ , the periodic action  $W_{p,q}^n = \sum_{j \in B_p} S_j^n$  is a Morse function on  $\mathbb{X}_{p,q}$ .
4. The gradients converge uniformly:  $\lim_{n \rightarrow \infty} \nabla S_j^n = \nabla S_j$  uniformly in  $C^1(\mathbb{R}^{\mathbb{Z}^d})$ .
5. The potentials converge uniformly on compacts:  $\lim_{n \rightarrow \infty} S_j^n = S_j$  uniformly on  $B_{p,q}$ .

In dimension  $d = 1$ , this theorem was proved by Golé [8,9] in the context of twist maps. His proof does not generalize to dimensions  $d > 1$  or to general monotone variational problems in dimension  $d = 1$ , because it explicitly exploits the interpretation of  $S_j(x) = S(x_j, x_{j+1})$  as the generating function of a twist map of the annulus, see Appendix A.

Our proof in higher dimensions is different, and it is based on Lemma 3.10 and ideas from equivariant Morse theory. We start by making Lemma 3.10 a bit more quantitative:

**Lemma 8.2.** Let  $\omega \in \mathbb{Q}^d$ . Then the  $\tau$ -action of  $(\mathbb{Z}^d \times \mathbb{Z})/I_\omega$  on  $\overline{\mathbb{X}}_\omega$  is properly discontinuous. More precisely, when  $(k, l)$  represents a nontrivial element of  $(\mathbb{Z}^d \times \mathbb{Z})/I_\omega$  and  $x \in \overline{\mathbb{X}}_\omega$ , then

$$|\tau_{k,l}x - x|_1 := \sum_{i \in B_p} |(\tau_{k,l}x - x)_i| \geq 1.$$

**Proof.** Let  $(p, q)$  be principal periods for  $\omega$ , i.e.  $\omega = -p^{-T}q$ , and write  $n := |\det p|$ . We notice that for an arbitrary  $k \in \mathbb{Z}^d$  it holds that  $-n\langle k, \omega \rangle = \langle k, |\det p| p^{-T}q \rangle \in \mathbb{Z}$  and thus that  $(nk, -n\langle \omega, k \rangle) \in I_\omega$ . Hence, writing  $nl = -n\langle \omega, k \rangle + nl + n\langle \omega, k \rangle$ , we see that  $\tau_{k,l}^n x = \tau_{nk, nl} x = \tau_{0, n(l + \langle \omega, k \rangle)} x$ . Thus,  $|\tau_{k,l}^n x - x|_1 = n^2 |l + \langle \omega, k \rangle|$ .

Now if  $(k, l)$  represents a nontrivial element of  $(\mathbb{Z}^d \times \mathbb{Z})/I_\omega$ , then  $n \cdot |l + \langle \omega, k \rangle| \geq 1$ , and hence we have that  $|\tau_{k,l}^n x - x|_1 \geq n$ . We claim that this implies that  $|\tau_{k,l}x - x|_1 \geq 1$ . This follows from the fact that  $\tau_{k,l}^{j+1}x - \tau_{k,l}^jx = \tau_{k,0}(\tau_{k,l}^jx - \tau_{k,l}^{j-1}x)$  and thus, by induction, that  $|\tau_{k,l}^{j+1}x - \tau_{k,l}^jx|_1 = |\tau_{k,l}x - x|_1$ . Therefore,  $|\tau_{k,l}^n x - x|_1 \leq |\tau_{k,l}^n x - \tau_{k,l}^{n-1}x|_1 + \dots + |\tau_{k,l}x - x|_1 = n|\tau_{k,l}x - x|_1$ , which means that  $|\tau_{k,l}x - x|_1 \geq 1$ .  $\square$

With Lemma 8.2 at hand, one can prove that the quotient  $\overline{\mathbb{X}}_\omega/(\mathbb{Z}^d \times \mathbb{Z})$  is a smooth manifold. An arbitrary  $\mathbb{Z}^d \times \mathbb{Z}$ -invariant function  $f : \overline{\mathbb{X}}_\omega \rightarrow \mathbb{R}$  descends to this quotient and can hence be perturbed into a shift-invariant Morse function  $f^\varepsilon$ . Instead of providing this rather standard construction from equivariant Morse theory, let us prove this latter fact directly here:

**Theorem 8.3.** Let  $\omega \in \mathbb{Q}^d$  and let  $p, q$  be principal periods for  $\omega$ . When  $f : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is an  $m \geq 2$  times continuously differentiable shift-invariant function, then for every  $\varepsilon > 0$  there exists a shift-invariant Morse function  $f^\varepsilon : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  with

$$\|f - f^\varepsilon\|_{C^m(E(N))} \leq \frac{\varepsilon}{(1 + N^2)^2} \quad \text{for every } N > 0.$$

Here,

$$E(N) := \{x \in \mathbb{X}_{p,q} \mid |x_i - x_k| \geq N \text{ for some } i \neq k \text{ with } i, k \in B_p\}.$$

**Proof.** Let  $c := \frac{1}{4|\det p|}$  and define the discrete collection of configurations

$$G_{p,q} := \{x : \mathbb{Z}^d \rightarrow c \cdot \mathbb{Z} \mid \tau_{p_j, q_j}x = x \text{ for } j = 1, \dots, d\} \subset \mathbb{X}_{p,q}.$$

We first of all remark that it is clear that  $\tau_{k,l}G_{p,q} = G_{p,q}$ . For  $x \in G_{p,q}$ , let us now define the balls

$$B_r(x) := \left\{ y \in \mathbb{X}_{p,q} \mid |y - x|_1 := \sum_{i \in B_p} |x_i - y_i| < r \right\}.$$

Then we have that  $\tau_{k,l}B_r(x) = B_r(\tau_{k,l}x)$ , because  $|\tau_{k,l}x - \tau_{k,l}y|_1 = |x - y|_1$ , that is the norm  $|\cdot|_1$  on  $\mathbb{X}_{p,q}$  is shift-invariant.

Moreover, if  $y \in \mathbb{X}_{p,q}$ , then there must be an element  $x \in G_{p,q}$  with  $|x_i - y_i| \leq \frac{1}{2}c$  for all  $i$ , that is for which  $|x - y|_1 \leq \frac{1}{8}$ . In other words,  $\mathbb{X}_{p,q} = \bigcup_{x \in G_{p,q}} B_r(x)$  when  $r > \frac{1}{8}$ . On the other hand, Lemma 8.2 implies that when  $r < \frac{1}{2}$ , then  $B_r(x) \cap B_r(\tau_{k,l}x) = \emptyset$  unless  $\langle \omega, k \rangle + l = 0$ .

This proves that for  $\frac{1}{8} < r < \frac{1}{2}$ , the collection  $\{B_r(x)\}_{x \in G_{p,q}}$  forms a shift-invariant covering of  $\mathbb{X}_{p,q}$  on which  $(\mathbb{Z}^d \times \mathbb{Z})/I_\omega$  acts “properly discontinuously”.

Finally, we let  $\phi : \mathbb{X}_{p,q} \rightarrow [0, 1]$  be a  $C^\infty$  bump function with the properties that  $\phi \equiv 0$  outside  $B_{\frac{1}{2}}(0)$  and  $\phi \equiv 1$  on  $B_{\frac{1}{4}}(0)$ . Let's say that  $\|\phi\|_{C^m(\mathbb{X}_{p,q})} \leq E$ .

After these preparations, we are ready to construct the perturbation  $f^\varepsilon$  of  $f$ . This is done by enumerating  $G_{p,q} = \{x^1, x^2, \dots\}$  and defining it inductively.

So let us assume that  $f_{n-1}^\varepsilon$  is  $\tau$ -invariant, satisfies the Morse property on the union  $\bigcup_{1 \leq i \leq n-1} B_{\frac{1}{4}}(x^i)$  and fulfills the estimates  $\|f - f_{n-1}^\varepsilon\|_{C^m(E(N))} \leq \frac{\varepsilon(1-2^{-(n-1)})}{(1+N^2)^2}$ .

We now want  $\alpha_n \in \mathbb{R}^{B_p}$  to be a vector so that  $x \mapsto f_{n-1}^\varepsilon(x) + \langle \alpha_n, x \rangle$  is Morse on  $B_{\frac{1}{4}}(x^n)$ . Such  $\alpha_n$ 's are dense in  $\mathbb{R}^{B_p}$  by Sard's theorem, see for instance [10].

The function  $f_n^\varepsilon$  is now defined as the shift-invariant function

$$f_n^\varepsilon(x) := f_{n-1}^\varepsilon(x) + \sum_{(k,l) \in (\mathbb{Z}^d \times \mathbb{Z})/I_\omega} \phi(\tau_{k,l}(x - x_n)) \langle \alpha_n, \tau_{k,l}(x - x^n) \rangle.$$

Because  $B_{\frac{1}{2}}(\tau_{k,l}x_n)$  does not intersect  $B_{\frac{1}{2}}(\tau_{K,L}x_n)$  unless  $(k,l) = (K,L) \bmod I_\omega$ , we have that at every  $x \in \mathbb{X}_{p,q}$ , the above sum consists of only one term. Moreover,  $f_n^\varepsilon$  is Morse on  $B_{\frac{1}{4}}(x^n)$  by construction.

In fact, by choosing  $\alpha_n$  small enough, one can make sure that  $f_n^\varepsilon$  is Morse on the entire union  $\bigcup_{1 \leq i \leq n} B_{\frac{1}{4}}(x^i)$ . This is true because the collection of Morse functions is open in the space of differentiable functions  $C^m(\bigcup_{1 \leq i \leq n-1} B_{\frac{1}{4}}(x^i))$  for  $m \geq 2$ , see [10].

By choosing  $\alpha_n$  even smaller if necessary, we can also arrange that  $f_{n-1}^\varepsilon - f_n^\varepsilon$  has a  $C^m(E(N))$ -norm less than  $\frac{2^{-n}\varepsilon}{(1+N^2)^2}$ . This implies that

$$\|f - f_n^\varepsilon\|_{C^m(E(N))} \leq \frac{(1-2^{-(n-1)})\varepsilon}{(1+N^2)^2} + \frac{2^{-n}\varepsilon}{(1+N^2)^2} = \frac{(1-2^{-n})\varepsilon}{(1+N^2)^2}.$$

The required  $f^\varepsilon$  is the limit  $f^\varepsilon := \lim_{n \rightarrow \infty} f_n^\varepsilon$ . Not only does this limit satisfy the required estimates, but it also stabilizes pointwise, which shows that it is Morse.  $\square$

We can now complete the proof of Theorem 8.1:

**Proof of Theorem 8.1.** We start by perturbing the  $S_j$  so that they satisfy a strict monotonicity criterion. This will then allow us to perturb the potentials once more without risking to destroy monotonicity condition D. Recall that  $B_p := p([0, 1]^d) \cap \mathbb{Z}^d$  is a fundamental domain of  $p$ . It has cardinality  $|\det p|$ . Our first perturbation step is now made by defining

$$\tilde{S}_j^n(x) := S_j(x) + \frac{1}{n} \sum_{i, k \in j+B_p} (x_k - x_i) \arctan(x_k - x_i).$$

The strict monotonicity of the  $\tilde{S}_j^n$  follows because  $\frac{d^2}{dx^2}(x \arctan x) = \frac{2}{(1+x^2)^2}$  is strictly positive. Hence,

$$\partial_{i,k} \tilde{S}_j^n(x) \leq -\frac{1}{n} \frac{2}{(1+(x_i - x_k)^2)^2} < 0 \quad \text{for all } i, k \in j+B_p \text{ with } i \neq k.$$

By Theorem 8.3, the periodic action  $\tilde{W}_{p,q}^n : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  defined by  $\tilde{W}_{p,q}^n(x) := \sum_{j \in B_p} \tilde{S}_j^n(x)$  can now be perturbed into a  $\tau$ -invariant Morse function  $W_{p,q}^n : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  of the form

$$W_{p,q}^n(x) = \tilde{W}_{p,q}^n(x) + F^n(x).$$

The perturbation  $F^n$  may be chosen so that it satisfies  $\|F^n\|_{C^2(\mathbb{X}_{p,q})} \leq \frac{1}{n}$ ,  $F^n(\tau_{k,l}x) = F^n(x)$  for all  $x \in \mathbb{X}_{p,q}$  and all  $k, l$  and  $|\partial_{i,k} F^n(x)| \leq \frac{1}{2n} \frac{1}{(1+(x_i - x_k)^2)^2}$ .

For  $x : \mathbb{Z}^d \rightarrow \mathbb{R}$ , let us denote by  $x|_{j+B_p}^{\text{per}}$  the  $p$ -periodic extension of  $x|_{j+B_p}$  defined by  $(x|_{j+B_p}^{\text{per}})_i = x_k$ , where  $k \in j + B_p$  is the unique element of  $j + B_p$  equal to  $i$  modulo  $p(\mathbb{Z}^d)$ . Then we can define, for each  $j \in \mathbb{Z}^d$  the new local potential

$$S_j^n(x) := \tilde{S}_j^n(x) + \frac{1}{|\det p|} F^n(x|_{j+B_p}^{\text{per}}).$$

We will now prove that these  $S_j^n$  satisfy all requirements of Theorem 8.1.

In fact, condition A and requirement 2 hold true because the range of interaction of both the sum  $\frac{1}{2n} \sum_{i,k \in j+B_p} (x_k - x_i) \arctan(x_k - x_i)$  and the perturbation  $F^n(x|_{j+B_p}^{\text{per}})$  do not exceed the bounded radius of  $B_p$ .

Condition B holds by definition. Condition C holds because  $x \mapsto x \arctan x$  is nonnegative and  $|F_n(x)|$  is uniformly bounded. Condition D holds true because  $\partial_{i,k} S_j^n(x) \leq \partial_{i,k} S_j(x)$ , as is easy to check. Requirement 3 holds because  $W_{p,q}^n(x) = \sum_{j \in B_p} S_j^n(x) = \tilde{W}_{p,q}^n(x) + F^n(x)$  is a Morse function by construction.

Requirement 4 and condition E are true because both  $|\frac{d}{dx} x \arctan x| = |\frac{1}{1+x^2} + \arctan x| \leq 3$  and  $|\frac{d^2}{dx^2} x \arctan x| = |\frac{2}{(1+x^2)^2}| \leq 2$  are uniformly bounded and  $\|F^n\|_{C^2(\mathbb{X}_{p,q})} \leq \frac{1}{2n}$ , so that  $\|\nabla S_j^n - \nabla S_j\|_{C^1(\mathbb{R}^{\mathbb{Z}^d})} \leq \frac{C}{n}$  for some constant  $C$  depending on the dimension  $d$ , the periodicity  $p$  and the range of interaction  $r$ .

Similarly, requirement 5 holds true because  $\sum_{i,k \in j+B_p} (x_i - x_k) \arctan(x_i - x_k)$  is uniformly bounded on  $B_{p,q}$  and  $|F_n(x)| \leq \frac{1}{2n}$  uniformly on  $\mathbb{X}_{p,q}$ .  $\square$

## 8.2. Existence of periodic ghost circles for Morse actions

We will now show that when the local potentials  $S_j$  satisfy conditions A–E and are chosen so that  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is a Morse function, then they admit a periodic ghost circle  $\Gamma \subset \mathbb{X}_{p,q}$ . More precisely, we will prove the following:

**Theorem 8.4.** *Let  $\omega \in \mathbb{Q}^d$  and let  $(p_1, q_1), \dots, (p_d, q_d)$  be principal periods for  $\omega$ . Assume moreover that the local potentials  $S_j$  are chosen so that  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is a Morse function. Then there exists a  $C^1$  ghost circle  $\Gamma \subset \mathbb{X}_{p,q}$  for the  $S_j$ . This ghost circle includes all the global minimizers of  $W_{p,q}$ . It consists of stationary points of index 0 and index 1 and heteroclinic orbits of the negative gradient flow.*

The construction of this ghost circle is essentially the same as the construction in dimension  $d = 1$  provided by Golé [8]. We nevertheless decided to provide the proofs.

To prove Theorem 8.4, we need two lemmas and the following definition:

**Definition 8.5.** We say that  $x \ll y$  are consecutive index-0 stationary configurations if there is no index-0 stationary configuration  $z$  with  $x \ll z \ll y$ .

It turns out that when  $W_{p,q}$  is a Morse function, then between consecutive index-0 stationary configurations we can find another critical point:

**Lemma 8.6** (Mountain Pass Theorem). *Assume that  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is Morse and let  $x \ll y$  be two consecutive index-0 stationary configurations of  $-\nabla W_{p,q}$ . Then there is an index-1 stationary configuration  $z$  in between  $x$  and  $y$ .*

**Proof.** We use a simple variant of the mountain pass theorem, see for instance [6, Section 8.5.1]. For this purpose, we let  $\mathcal{C}$  be the collection of curves from  $x$  to  $y$  lying in the order interval  $K := [x, y]$ ,



that is

$$C := \{\gamma : [0, 1] \rightarrow K \mid \gamma(0) = x, \gamma(1) = y \text{ and } \gamma \text{ is continuous}\}.$$

We now claim that there is a critical point  $z \in \hat{K}$  for which  $W_{p,q}(z) = c$ , where

$$c := \inf_{\gamma \in C} \max_{0 \leq t \leq 1} W_{p,q}(\gamma(t)).$$

To prove our claim, let us define, for  $\delta \in \mathbb{R}$ , the sub-level sets

$$K^\delta := \{x \in K \mid W_{p,q}(x) \leq \delta\}.$$

These  $K^\delta$  are invariant under the forward flow of  $\frac{dx}{dt} = -\nabla W_{p,q}(x)$ . This is true because  $K$  is invariant and because  $W_{p,q}$  is a Lyapunov function for the gradient flow.

Suppose now that there is no critical point  $x \ll z \ll y$  with  $W_{p,q}(z) = c$ . We will show that this leads to a contradiction.

We first of all remark that, by the Morse lemma and the fact that  $x$  and  $y$  have index 0, it holds that  $c > \max\{W_{p,q}(x), W_{p,q}(y)\}$ . Thus, because there are only finitely many critical points in  $K$ , and none of these except  $x$  and  $y$  lie in  $\partial K$ , there exists an  $\varepsilon > 0$  so that the set  $K^{c+\varepsilon} \setminus K^{c-\varepsilon}$  does not contain any critical points.

This in turn implies, by compactness, that there is a  $\sigma > 0$  so that  $\|\nabla W_{p,q}\|^2 > \sigma$  on  $K^{c+\varepsilon} \setminus K^{c-\varepsilon/2}$ . Hence, a solution curve  $t \mapsto x(t)$  of the negative gradient flow satisfies  $\frac{d}{dt} W_{p,q}(x(t)) = -\|\nabla W_{p,q}(x(t))\|^2 < -\sigma$  so long as  $x(t) \in K^{c+\varepsilon} \setminus K^{c-\varepsilon/2}$ . In particular, there is a  $T > 0$  for which  $\psi_T(K^{c+\varepsilon}) \subset K^{c-\varepsilon/2}$ .

At the same time, by definition of  $c$ , there exists a  $\gamma \in C$  with  $\gamma([0, 1]) \subset K^{c+\varepsilon}$ . The curve  $\psi_T \circ \gamma \in C$  then lies entirely in  $K^{c-\varepsilon/2}$ . This contradicts the definition of  $c$  and hence there must be critical points  $x \ll z_1, \dots, z_m \ll y$  with  $W_{p,q}(z_i) = c$ .

It remains to show that at least one of the  $z_i$  has index one. In fact, the argument is a bit subtle. We start by observing the following:

1. If  $x \ll z_i \ll y$  is an index-0 critical point with  $W_{p,q}(z_i) = c$ , then there are  $\alpha_i, \beta_i > 0$  so that whenever  $\gamma \in C$  intersects  $B_{\alpha_i}(z_i)$ , then  $\max_{t \in [0, 1]} W_{p,q}(\gamma(t)) \geq c + \beta_i$ .
2. If  $x \ll z_i \ll y$  is an index- $\geq 2$  critical point with  $W_{p,q}(z_i) = c$ , then there is an  $\alpha_i > 0$  so that whenever  $\gamma \in C$  intersects  $B_{\alpha_i}(z_i)$ , then  $\gamma$  is homotopic to a curve  $\tilde{\gamma} \in C$  with the property that  $\tilde{\gamma}$  does not intersect  $B_{\alpha_i}(z_i)$ , while  $\max_{t \in [0, 1]} W_{p,q}(\tilde{\gamma}(t)) \leq \max_{t \in [0, 1]} W_{p,q}(\gamma(t))$ .

These statements are easy to prove in local Morse coordinates near the critical point  $z_i$ . At the same time, by compactness, we have that there exist  $\varepsilon, \sigma_1, \sigma_2 > 0$  so that  $\sigma_1 < \|\nabla W_{p,q}\|^2 < \sigma_2$  on  $K^{c+\varepsilon} \setminus (K^{c-\varepsilon/2} \cup B_{\beta_1/2}(z_1) \cup \dots \cup B_{\beta_m/2}(z_m))$ . Using that  $\|\frac{dx(t)}{dt}\| = \|\nabla W_{p,q}(x(t))\|$  and  $\frac{d}{dt} W_{p,q}(x(t)) = -\|\nabla W_{p,q}(x(t))\|^2$  for solutions of the gradient flow, one can prove quite easily that this implies:

3. If  $t \mapsto x(t)$  solves  $\frac{dx}{dt} = -\nabla W_{p,q}(x)$  and  $W_{p,q}(x(0)) \leq c + \varepsilon$  and  $x(0) \notin B_{\beta_1}(z_1) \cup \dots \cup B_{\beta_m}(z_m)$ , then for  $0 \leq t \leq T := \min\{\beta_1, \dots, \beta_m\}/2\sqrt{\sigma_2}$  one has that  $x(t) \notin \bigcup_{i=1}^m B_{\beta_i/2}(z_i)$  and hence  $W_{p,q}(x(T)) \leq \max\{c - \varepsilon/2, W_{p,q}(x(0)) - T\sigma_1\}$ .

We now use these facts as follows: Let us assume that none of the  $z_i$  has index 1 and let  $\gamma_n \in C$  be a sequence of curves with  $c \leq \max_{t \in [0, 1]} W_{p,q}(\gamma_n(t)) \leq c + \frac{1}{n}$ . By property 1 we know that for large enough  $n$ , the curve  $\gamma_n$  does not intersect  $B_{\alpha_i}(z_i)$  for any of the index-0 points  $z_i$ . At the same time, by property 2 we may assume that none of the  $\gamma_n$  intersects the  $B_{\alpha_i}(z_i)$  for any of the index- $\geq 2$  points  $z_i$ . Property 3 then implies that for large enough  $n$  we have that  $(\psi_T \circ \gamma_n)([0, 1]) \subset K^{c-\delta}$  for some  $\delta > 0$ . This contradicts the definition of  $c$ .  $\square$

The next step is to show that the unstable manifold of the index-1 critical point  $z$  of Lemma 8.6 defines  $C^1$  ordered heteroclinic connections to its neighboring index-0 critical points  $x$  and  $y$ . This result, in more generality, can also be found in [1, Theorem 1 of Chapter 4].

**Lemma 8.7.** *Let  $W_{p,q} : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  be Morse, let  $x \ll y$  be two consecutive index-0 stationary configurations of  $W_{p,q}$  and let  $z$  be an index-1 stationary configuration with  $x \ll z \ll y$ . Then the unstable manifold of  $z$  forms strictly ordered heteroclinic connections from  $z$  to  $x$  and  $y$ .*

**Proof.** We consider the linearization of the negative gradient vector field at  $z$ , given by the matrix  $-D^2W_{p,q}(z)$ . The twist condition D and the bound on the second derivatives E together guarantee that there exists a constant  $M > 0$  so that the symmetric matrix  $H := -D^2W_{p,q}(z) + M \text{Id}$  is nonnegative and strictly positive on its diagonal and its two off-diagonals.

By the theorem of Perron–Frobenius,  $H$  then has to have a unique simple largest eigenvalue  $\lambda_{\max} + M \in \mathbb{R}_+$  and the corresponding eigenvector  $e_{\max}$  can be chosen strictly positive. Because  $z$  is an index-1 point,  $\lambda_{\max}$  is then the unique positive eigenvalue of  $-D^2W_{p,q}(z)$  and  $e_{\max}$  is its strictly positive eigenvector.

The unstable manifold  $\mathcal{W}^u(z)$  of  $z$  is one-dimensional and at  $z$  it is tangent to  $e_{\max}$ . In fact, it consists of  $z$  and two orbits of the negative gradient flow

$$\alpha^\pm(t) = z \pm e^{\lambda_{\max} t} e_{\max} + o(e^{\lambda_{\max} t}) \quad \text{for } t \rightarrow -\infty.$$

In particular we see that close to  $z$ , the unstable manifold is strictly ordered, because  $e_{\max}$  is strictly positive. Theorem 6.2 then implies that the entire  $\mathcal{W}^u(z)$  is strictly ordered. Thus, we see that there must be two critical points  $z^- := \lim_{t \rightarrow \infty} \alpha_-(t)$  and  $z^+ := \lim_{t \rightarrow \infty} \alpha_+(t)$ . We claim that  $z^- = x$  and  $z^+ = y$ .

To prove this, we will show that  $z^-$  and  $z^+$  are index-0 critical points. Our claim then follows because  $x \leq z^- \ll z \ll z^+ \leq y$  by monotonicity and because  $x$  and  $y$  are consecutive index-0 points.

So let us consider the linearization matrix  $-D^2W_{p,q}(z^-)$ . It also has a unique maximal eigenvalue  $\lambda_{\max}^-$  and positive eigenvector  $e_{\max}^-$ . We know that  $\lim_{t \rightarrow \infty} \alpha^-(t) = z^-$  and that  $\alpha^-(\mathbb{R})$  is strictly ordered. At the same time, because  $-D^2W_{p,q}(z^-)$  is symmetric,  $e_{\max}^-$  is perpendicular to all other eigenvectors of  $-D^2W_{p,q}(z^-)$ , which implies that none of these other eigenvectors lies in the positive or the negative quadrant. This means that  $\alpha^-(t)$  has to approach  $z^-$  tangent to  $e_{\max}^-$ , that is

$$\alpha^-(t) = z^- + e^{\lambda_{\max}^- t} e_{\max}^- + o(e^{\lambda_{\max}^- t}) \quad \text{for } t \rightarrow \infty.$$

In particular,  $\lambda_{\max}^- < 0$ . But  $\lambda_{\max}^-$  is the maximal eigenvalue of  $-D^2W_{p,q}(z^-)$ . This means that all eigenvalues of  $D^2W_{p,q}(z^-)$  are positive, i.e. that  $z^-$  is an index-0 point.

A similar argument for  $z^+$  finishes the proof.  $\square$

We conclude with a definition and then give the proof of Theorem 8.4.

**Definition 8.8.** A nonempty and strictly ordered collection of configurations

$$C_0 = \{\cdots \ll x_{-1} \ll x_0 \ll x_1 \ll \cdots\} \subset \mathbb{X}_{p,q}$$

is called a *maximal index-0 skeleton* for the Morse function  $W_{p,q}$  if:

- it consists of index-0 critical points of  $W_{p,q}$ ;
- it is shift-invariant: for all  $x \in C_0$  and  $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}$ , it holds that  $\tau_{k,l} x \in C_0$ ;
- it is maximal: if  $y \notin C_0$  is an index-0 point, then there is no  $i \in \mathbb{Z}$  with  $x_i \ll y \ll x_{i+1}$ .

**Proof of Theorem 8.4.** We remark that a maximal index-0 skeleton in general is not unique, but it is not hard to see that a maximal index-0 skeleton exists if  $W_{p,q}$  is Morse.

Indeed, one can construct one by starting with the strictly ordered, shift-invariant collection  $C_0^0 = \{\dots, \bar{x}_{-1}, \bar{x}_0, \bar{x}_1, \dots\}$  of all the global minimizers of  $W_{p,q}$ . We note that  $C_0^0$  is discrete because  $W_{p,q}$  is a Morse function.

If there exists an index-0 point  $x \notin C_0^0$  with the property that  $\bar{x}_i \ll x \ll \bar{x}_i$  for some  $i$ , then one augments  $C_0^0$  by the  $\tau$ -orbit of this  $x$ , thus obtaining the strictly ordered, shift-invariant and discrete collection

$$C_0^1 := C_0^0 \cup \{\tau_{k,l}x \mid (k, l) \in \mathbb{Z}^d \times \mathbb{Z}\}.$$

One keeps on adding  $\tau$ -orbits of index-0 points this way. The Morse property of  $W_{p,q}$  guarantees that the number of index-0 points between  $\bar{x}_0$  and  $\bar{x}_0 + 1$  is finite, which implies that this process stops after finitely many steps.

The maximality of an index-0 maximal skeleton  $C_0 = \{\dots \ll x_{-1} \ll x_0 \ll x_1 \ll \dots\}$  just means that the pairs  $x_i, x_{i+1}$  are consecutive index-0 points. The Mountain Pass Theorem (Lemma 8.6) guarantees that between these consecutive elements, there is an index-1 critical point  $z_i$ , while Lemma 8.7 says that the unstable manifold of this  $z_i$  defines strictly ordered heteroclinic connections from  $z_i$  to  $x_i$  and  $x_{i+1}$ .

If we choose the  $z_i$  in such a way that  $\bar{C}_0 := \{\dots \ll x_{-1} \ll z_{-1} \ll x_0 \ll z_0 \ll x_1 \ll z_1 \ll \dots\}$  is shift-invariant, then the union of  $\bar{C}_0$  and these heteroclinic connections is a ghost circle  $\Gamma$ . The construction above shows that  $\Gamma$  may be assumed to contain all global minimizers of  $W_{p,q}$ .

It only remains to show that this  $\Gamma$  is  $C^1$ . This is clear except at the critical points. But in the proof of the Lemma 8.7, we have seen that at the critical points, the heteroclinic connections are tangent to the dominant eigenvector. This eigenvector is simple and hence,  $\Gamma$  is  $C^1$  also at critical points.  $\square$

## 9. Convergence of ghost circles

Section 8 was devoted to the construction of periodic ghost circles for action functions that satisfy the Morse property. In this section we will prove the existence of periodic ghost circles for arbitrary action functions. In turn, this will then imply the existence of ghost circles with irrational rotation vectors. These results follow from a compactness theorem for ghost circles that we will prove below. Before we can formulate it, let us specify what it means for a sequence of ghost circles to converge:

**Definition 9.1** (Convergence of ghost circles). We say that a sequence of ghost circles  $\Gamma_n$  converges to a ghost circle  $\Gamma_\infty$ , if for every  $\xi \in \mathbb{R}$ , the sequence of configurations  $x^n(\xi) \in \Gamma_n$  defined by  $\pi_0(x^n(\xi)) = \xi$  converges pointwise to the configuration  $x^\infty(\xi) \in \Gamma_\infty$  defined by  $\pi_0(x^\infty(\xi)) = \xi$ .

Thus, if  $\Gamma_n \rightarrow \Gamma_\infty$  as  $n \rightarrow \infty$  then  $\Gamma_\infty$  consists of pointwise limits of elements of the  $\Gamma_n$ .

Before stating the most important results of this section, let us make a few simple observations concerning convergence of ghost circles. First of all, one can observe that if  $\lim_{n \rightarrow \infty} \omega_n = \omega_\infty$  and if a sequence of ghost circles  $\Gamma_n \subset \mathcal{B}_{\omega_n}$  converges to  $\Gamma_\infty$ , then it must be true that  $\Gamma_\infty \subset \mathcal{B}_{\omega_\infty}$ . This follows from the continuity of the rotation vector as a function on  $\mathcal{B}$  and the fact that the rotation vector of a ghost circle is defined as the rotation vector of any of its elements.

The second remark is that if the  $\Gamma_n \subset \bar{\mathcal{B}}_\omega$  are periodic ghost circles with the same rational rotation vector, and  $\Gamma_n \rightarrow \Gamma_\infty$ , then  $\Gamma_\infty \subset \bar{\mathcal{B}}_\omega$  is periodic as well. This follows because  $\bar{\mathcal{B}}_\omega$  is a closed subset of  $\mathbb{R}^{\mathbb{Z}^d}$ .

Our compactness result now is the following:

**Theorem 9.2.** Let  $\omega_n \in K \subset \mathbb{R}^d$  be a sequence of rotation vectors contained in a compact set  $K$  and converging to  $\omega_\infty \in K \subset \mathbb{R}^d$  and let  $S_j^n : \mathbb{R}^d \rightarrow \mathbb{R}$  be a sequence of local potentials such that  $\nabla S_j^n$  converge to  $\nabla S_j^\infty$  uniformly in  $C^1(\mathbb{R}^d)$ . Finally, let  $\Gamma_n$  be a sequence of ghost circles for the  $S_j^n$  of rotation vector  $\omega_n$ .

Then there exists a ghost circle  $\Gamma_\infty$  for the  $S_j^\infty$  of rotation vector  $\omega_\infty$  and a subsequence  $\{n_j\}_{j=1}^\infty$  such that  $\lim_{j \rightarrow \infty} \Gamma_{n_j} = \Gamma_\infty$ .

If moreover  $\lim_{n \rightarrow \infty} S_j^n = S_j^\infty$  uniformly in  $C^0(\mathcal{B}_K)$  and if every  $\Gamma_n$  contains a global minimizer, then also  $\Gamma_\infty$  contains a global minimizer.

Before proving this compactness result, let us formulate its two main implications:

**Theorem 9.3.** Let  $\omega \in \mathbb{Q}^d$  be arbitrary and let the local potentials  $S_j$  be given. Then there exists a periodic ghost circle  $\Gamma_\omega \subset \bar{\mathcal{B}}_\omega$  for the  $S_j$ . This  $\Gamma_\omega$  may be chosen so that it contains a global minimizer.

**Proof.** Given  $\omega \in \mathbb{Q}^d$  and any local potentials  $S_j$ , choose principal periods  $(p, q)$  for  $\omega$ . By Theorem 8.1 we can choose a sequence of local potentials  $S_j^n$  such that  $\lim_{n \rightarrow \infty} \nabla S_j^n = \nabla S_j$  uniformly in  $C^1(\mathbb{R}^{\mathbb{Z}^d})$  and  $\lim_{n \rightarrow \infty} S_j^n = S_j$  uniformly in  $C^0(\mathcal{B}_K)$ , while at the same time  $W_{p,q}^n : \mathbb{X}_{p,q} \rightarrow \mathbb{R}$  is a Morse function. Then, by Theorem 8.4, there is a ghost circle  $\Gamma_n \subset \mathcal{B}_{p,q} = \bar{\mathcal{B}}_\omega$  for the local potentials  $S_j^n$  that contains a minimizer of  $W_{p,q}^n$ . By Theorem 9.2, a subsequence of the  $\Gamma_n$  converges to a ghost circle  $\Gamma_\omega \subset \mathcal{B}_{p,q} = \bar{\mathcal{B}}_\omega$  for the local potentials  $S_j$ . By the second conclusion of Theorem 9.2,  $\Gamma_\omega$  contains a global minimizer.  $\square$

**Theorem 9.4.** Let  $\omega \in \mathbb{R}^d \setminus \mathbb{Q}^d$  and let the local potentials  $S_j$  be given. Then there exists a ghost circle  $\Gamma_\omega \subset \bar{\mathcal{B}}_\omega$  for the  $S_j$ . This  $\Gamma_\omega$  may be chosen so that it contains the entire Aubry–Mather set of rotation vector  $\omega$ .

**Proof.** Given  $\omega \in \mathbb{R}^d$  and local potentials  $S_j$ , choose a sequence  $\omega_n \in \mathbb{Q}^d$  such that  $\lim_{n \rightarrow \infty} \omega_n = \omega$ . By Theorem 9.3, there is a periodic ghost circle  $\Gamma_n \subset \bar{\mathcal{B}}_{\omega_n}$  for the local potentials  $S_j$  that contains at least one global minimizer. By Theorem 9.2, a subsequence of the  $\Gamma_n$  converges to a ghost circle  $\Gamma_\omega \subset \bar{\mathcal{B}}_\omega$ .

The requirement for the second conclusion of Theorem 9.2 is trivially valid, so that  $\Gamma_\omega$  contains a global minimizer, say  $x$ . Being closed and shift-invariant, this implies that  $\Gamma_\omega$  contains the entire Aubry–Mather set  $\mathcal{M}(x)$ .  $\square$

Before proving Theorem 9.2, we remark that if  $\Gamma_n$  is an arbitrary sequence of ghost circles for the local potentials  $S_j^n$  and with rotation vectors  $\omega_n$  in a compact set  $K$ , then for every  $\xi \in \mathbb{R}$  the sequence of configurations  $x^n(\xi) \in \Gamma_n$  has a subsequence that converges pointwise. This just follows from the compactness of  $\mathcal{B}_K \cap \{x \in \mathbb{R}^{\mathbb{Z}^d} \mid \pi_0(x) = \xi\}$ . The problem is to show that this subsequence can be chosen independent of  $\xi$  and that the collection of limit configurations  $\{\lim_{n \rightarrow \infty} x^n(\xi) \mid \xi \in \mathbb{R}\}$  forms a ghost circle for the  $S_j = \lim_{n \rightarrow \infty} S_j^n$ .

We will now make some preparations for the proof of Theorem 9.2. To start with, we define for a given ghost circle  $\Gamma$ , the map

$$T^\Gamma : \mathbb{R} \rightarrow \Gamma \text{ by } T^\Gamma := \Psi_{-1} \circ (\pi_0|_\Gamma)^{-1}, \text{ that is: } T_k^\Gamma(\xi) := (T^\Gamma(\xi))_k = (\pi_k \circ \Psi_{-1} \circ (\pi_0|_\Gamma)^{-1})(\xi).$$

Here,  $\Psi_{-1} : \mathbb{X} \rightarrow \mathbb{X}$  denotes the time-(-1) flow of  $\frac{dx}{dt} = -\nabla W(x)$ . By Theorem 5.4,  $T^\Gamma$  is a homeomorphism, being the composition of two homeomorphisms. Moreover, it is “pointwise Lipschitz continuous”:

**Lemma 9.5.** Let  $K \subset \mathbb{R}^d$  be a compact set and  $\Gamma \subset \mathcal{B}_K = \bigcup_{\omega \in K} \mathcal{B}_\omega$  a ghost circle with rotation vector  $\omega \in K$  for the local potentials  $S_j$  satisfying conditions A–E. Then, for every  $k \in \mathbb{Z}^d$ , there is a constant  $A_{\|k\|} > 0$ , depending only on  $K$  and  $\|k\|$  such that

$$|T_k^\Gamma(\xi) - T_k^\Gamma(v)| \leq A_{\|k\|} |\xi - v|.$$

**Proof.** Let  $\xi, \nu \in \mathbb{R}$  and denote  $X = (\pi_0|_\Gamma)^{-1}(\xi)$  and  $Y = (\pi_0|_\Gamma)^{-1}(\nu)$ . Assume that  $\xi < \nu$ , whence  $X \ll Y$ . Denote by  $\Psi_t$  the time- $t$  flow of  $-\nabla W$ , with  $W := \sum_{j \in \mathbb{Z}^d} S_j$ . Since  $\Gamma$  is forward and backward invariant under  $\Psi$ , both  $\Psi_{-1}(X)$  and  $\Psi_{-1}(Y)$  lie in  $\Gamma$  and satisfy  $\Psi_{-1}(X) \ll \Psi_{-1}(Y)$ . Now we apply the parabolic Harnack inequality of Theorem 6.4 to  $t = 1$ ,  $i = 0$ ,  $x = \Psi_{-1}(X)$  and  $y = \Psi_{-1}(Y)$ , to find that there is an  $L > 0$  depending only on  $K$  and  $\|k\|$  such that

$$T_k^\Gamma(\nu) - T_k^\Gamma(\xi) = (\Psi_{-1}Y)_k - (\Psi_{-1}X)_k \leq \frac{1}{L}(Y_0 - X_0) = \frac{1}{L}(\nu - \xi).$$

A similar argument in the case that  $\xi > \nu$  finishes the proof.  $\square$

We remark here that we see no reason why the maps  $\pi_k \circ (\pi_0|_\Gamma)^{-1}$  should be uniformly Lipschitz continuous. This is why we study the maps  $\pi_k \circ \Psi_{-1} \circ (\pi_0|_\Gamma)^{-1}$  instead.

**Definition 9.6.** We say that a sequence of maps  $T^n : \mathbb{R} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$  converges *pointwise uniformly* to a map  $T^\infty : \mathbb{R} \rightarrow \mathbb{R}^{\mathbb{Z}^d}$  as  $n \rightarrow \infty$  if for every  $k \in \mathbb{Z}^d$  the sequence of maps  $T_k^n := \pi_k \circ T^n : \mathbb{R} \rightarrow \mathbb{R}$  converges uniformly to  $T_k^\infty := \pi_k \circ T^\infty$  as  $n \rightarrow \infty$ .

**Corollary 9.7.** Let  $K \subset \mathbb{R}^d$  be a compact set. Assume that for every  $n \in \mathbb{N}$ , we are given a rotation vector  $\omega_n \in K$ , local potentials  $S_j^n$  satisfying conditions A–E and ghost circles  $\Gamma_n \subset \mathcal{B}_{\omega_n}$  for the local potentials  $S_j^n$ .

Then there is a subsequence  $\{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N}$  with the property that the maps  $T^{n_j} : \mathbb{R} \rightarrow \mathbb{R}$  converge pointwise uniformly on  $\mathbb{R}$ , say  $T^{n_j} \rightarrow T^\infty$  as  $j \rightarrow \infty$ . Each limit map  $T_k^\infty := \pi_k \circ T^\infty : \mathbb{R} \rightarrow \mathbb{R}$ , is nondecreasing, surjective and Lipschitz continuous.

**Proof.** Fix a  $k \in \mathbb{Z}^d$ . By Lemma 9.5, the maps  $T_k^{n_j} := \pi_k \circ T^{n_j}$  are uniformly-in- $n$  Lipschitz continuous with Lipschitz constant  $\Lambda_{\|k\|}$ . Moreover, by the definition of a ghost circle they are 1-periodic and increasing. To see that they are uniformly bounded on compacts, we then just have to note that Proposition 7.3 implies that  $T_0^{n_j}(0) \in [-1, 1]$ , while Lemma 3.5 and the fact that  $(\Psi_{-1} \circ (\pi_0^n)^{-1})(\xi)$  is a Birkhoff sequence then imply that  $T_k^{n_j}(0) \in [-2 - \|K\| \cdot \|k\|, 2 + \|K\| \cdot \|k\|]$ , where  $\|K\| := \max_{\omega \in K} \|\omega\|$ .

Thus, the theorem of Arzelà–Ascoli guarantees that there exists a uniformly convergent subsequence  $T_k^{n_{j,k}} \rightarrow T_k^\infty$  for  $j \rightarrow \infty$ . Clearly,  $T_k^\infty$  is nondecreasing and Lipschitz continuous with Lipschitz constant  $\Lambda_{\|k\|}$ .

Let  $j \mapsto k_j, \mathbb{N} \rightarrow \mathbb{Z}^d$  be a denumeration of  $\mathbb{Z}^d$ . Then the diagonal sequence  $\{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N}$  defined by  $n_j := n_{j,k_j}$  has the property that  $T^{n_j} \rightarrow T^\infty$  pointwise uniformly as  $j \rightarrow \infty$ .  $\square$

**Theorem 9.8** (Convergence of ghost circles). Let  $\Gamma_n$  be a sequence of ghost circles for the local potentials  $S_j^n$ . Assume that there are local potential functions  $S_j^\infty$  such that  $\nabla S_j^n \rightarrow \nabla S_j^\infty$  uniformly in  $C^1(\mathbb{R}^{\mathbb{Z}^d})$  and that the maps  $T^{n_j}$  converge pointwise uniformly. Then there is a ghost circle  $\Gamma_\infty$  for the local potentials  $S_j^\infty$  such that  $\Gamma_n \rightarrow \Gamma_\infty$  as  $n \rightarrow \infty$ .

Moreover, when  $\Gamma_n$  contains a global minimizer  $x_n$  and  $\lim_{n \rightarrow \infty} S_j^n = S_j^\infty$  uniformly in  $C^0(\mathcal{B}_K)$ , then  $\Gamma_\infty$  contains a global minimizer as well.

Theorem 9.2 now follows directly from Corollary 9.7 and Theorem 9.8.

Before we prove Theorem 9.8, let us recall that if  $\Gamma_\infty = \lim_{n \rightarrow \infty} \Gamma_n$  exists, then it must be equal to

$$\Gamma_\infty := \left\{ x^\infty(\xi) := \lim_{n \rightarrow \infty} x^n(\xi) \text{ pointwise} \mid \xi \in \mathbb{R} \right\}.$$

At this point, it is of course not clear whether the limit  $\lim_{n \rightarrow \infty} x^n(\xi)$  exists for every  $\xi \in \mathbb{R}$ . To see that it does under the conditions of Theorem 9.8, we note that  $x^n(\xi) = \Psi_1^n(T^n(\xi))$ , so that

$$\lim_{n \rightarrow \infty} x^n(\xi) = \lim_{n \rightarrow \infty} \Psi_1^n(T^n(\xi)) \text{ exists and is equal to } \Psi_1^\infty(T^\infty(\xi)).$$

This is true because on the one hand, according to Corollary 5.9,  $\Psi_1^n \rightarrow \Psi_1^\infty$  uniformly in the topology of pointwise convergence, while on the other hand it holds that for every  $\xi \in \mathbb{R}$ , the sequence of configurations  $T^{T^n}(\xi) \in \mathcal{B}$  converges pointwise to the configuration  $T^\infty(\xi)$  as  $n \rightarrow \infty$ , because  $T^{T^n} \rightarrow T^\infty$  pointwise uniformly. Thus we find that under the conditions of Theorem 9.8,  $\Gamma_\infty := \lim_{n \rightarrow \infty} \Gamma_n \subset \mathcal{B}_{\omega_\infty}$  is well defined and moreover that, if  $\Gamma_\infty$  is a ghost circle, then  $T^{\Gamma_\infty} = T^\infty$ .

We will now show that  $\Gamma_\infty$  is in fact a ghost circle for the local potentials  $S_j^\infty$ :

**Proof of Theorem 9.8.** We first check that  $\Gamma_\infty$  has the properties required for a ghost circle:

1. *Closedness*: Let  $x^\infty(\xi_m) \in \Gamma_\infty$  be a sequence of configurations that converges pointwise. This implies that the  $\xi_m$  converge, say to  $\xi$ . We now want to show that  $\lim_{m \rightarrow \infty} x^\infty(\xi_m) = x^\infty(\xi) \in \Gamma_\infty$  pointwise. This follows because  $\lim_{m \rightarrow \infty} x^\infty(\xi_m) = \lim_{m \rightarrow \infty} \Psi_1^\infty(T^\infty(\xi_m)) = \Psi_1^\infty(\lim_{m \rightarrow \infty} T^\infty(\xi_m)) = \Psi_1^\infty(T^\infty(\xi)) = x^\infty(\xi)$ . All these limits are pointwise. We have used that  $\Psi_1^\infty$  is continuous for pointwise convergence and that  $\lim_{m \rightarrow \infty} T^\infty(\xi_m) = T^\infty(\xi)$  pointwise.

2. *Connectedness*: We note that  $\Gamma_\infty = \Psi_1^\infty(T^\infty(\mathbb{R}))$ , so it is the image under a continuous map of a connected set, hence connected.

3. *Strict ordering*: Suppose  $\xi < v$ . Recall that  $T^\infty$  is nondecreasing, so  $T^\infty(\xi) \leq T^\infty(v)$ . We remark that  $T^\infty(\xi)$  cannot equal  $T^\infty(v)$ , because this would imply that  $\xi = \pi_0(x^\infty(\xi)) = \pi_0(\Psi_1^\infty(T^\infty(\xi))) = \pi_0(\Psi_1^\infty(T^\infty(v))) = \pi_0(x^\infty(v)) = v$ . Thus  $T^\infty(\xi) < T^\infty(v)$ .

The strict monotonicity of the negative gradient flow, then implies that  $x^\infty(\xi) = \Psi_1^\infty(T^\infty(\xi)) \ll \Psi_1^\infty(T^\infty(v)) = x^\infty(v)$ .

4. *Shift-invariance*: Let  $x^\infty \in \Gamma_\infty$ , that is  $x^\infty = \lim_{n \rightarrow \infty} x^n$  with  $x^n \in \Gamma_n$  and  $\pi_0(x^n) = \pi_0(x^\infty)$ . Let  $k \in \mathbb{Z}^d$  and  $l \in \mathbb{Z}$  be given. We want to show that  $\tau_{k,l}x^\infty \in \Gamma_\infty$ , that is we want to show that  $\tau_{k,l}x^\infty = \lim_{n \rightarrow \infty} y^n$  with  $y^n \in \Gamma_n$  such that  $\pi_0(y^n) = \pi_0(\tau_{k,l}x^\infty)$ . We prove this by writing

$$\lim_{n \rightarrow \infty} (\tau_{k,l}x^\infty - y^n) = \lim_{n \rightarrow \infty} (\tau_{k,l}x^\infty - \tau_{k,l}x^n) + \lim_{n \rightarrow \infty} (\tau_{k,l}x^n - y^n)$$

and showing that both limits on the right-hand side vanish.

The first limit is zero because, by Lemma 5.3,  $\tau_{k,l}$  is continuous in the topology of pointwise convergence. Thus we have that  $\lim_{n \rightarrow \infty} (\tau_{k,l}x^\infty - \tau_{k,l}x^n) = \lim_{n \rightarrow \infty} \tau_{k,0}(x^\infty - x^n) = 0$ .

For the second limit, we realize that  $\tau_{k,l}x^n \in \Gamma_n$  because  $\Gamma_n$  is shift-invariant and we observe that  $\pi_0(\tau_{k,l}x^n) = x_k^n + l$ . Because  $T^{T^n} \rightarrow T^\infty$  pointwise uniformly, we moreover know that  $\lim_{n \rightarrow \infty} T^{T^n}(x_k^n + l) = T^\infty(x_k^\infty + l)$  pointwise. Thus, by the uniform convergence of the  $\Psi_1^n$  to  $\Psi_1^\infty$ ,

$$\lim_{n \rightarrow \infty} (\tau_{k,l}x^n) = \lim_{n \rightarrow \infty} \Psi_1^n(T^n(x_k^n + l)) = \Psi_1^\infty(T^\infty(x_k^\infty + l)) = \lim_{n \rightarrow \infty} \Psi_1^\infty(T^n(x_k^\infty + l)) = \lim_{n \rightarrow \infty} y^n.$$

5. *Flow-invariance*: This is proved in a similar way as shift-invariance. So, let  $x^\infty \in \Gamma_\infty$ , that is  $x^\infty = \lim_{n \rightarrow \infty} x^n$  with  $x^n \in \Gamma_n$  and  $\pi_0(x^n) = \pi_0(x^\infty)$ . Let  $t \in \mathbb{R}$  be given. We want to show that  $\Psi_t^\infty x^\infty \in \Gamma_\infty$ , that is we want to show that  $\Psi_t^\infty x^\infty = \lim_{n \rightarrow \infty} y^n$  with  $y^n \in \Gamma_n$  such that  $\pi_0(y^n) = \pi_0(\Psi_t^\infty x^\infty)$ . We prove this by writing

$$\lim_{n \rightarrow \infty} (\Psi_t^\infty x^\infty - y^n) = \lim_{n \rightarrow \infty} (\Psi_t^\infty x^\infty - \Psi_t^n x^n) + \lim_{n \rightarrow \infty} (\Psi_t^n x^n - y^n)$$

and showing that both limits on the right-hand side vanish.

The first limit is zero because, by Theorem 5.4,  $\Psi_t^n$  converges to  $\Psi_t^\infty$  uniformly. Thus we have that  $\lim_{n \rightarrow \infty} \Psi_t^n x^n = \Psi_t^\infty x^\infty$ .

For the second limit, we realize that  $\Psi_t^n x^n \in \Gamma^n$  because  $\Gamma_n$  is flow-invariant and we observe that  $\lim_{n \rightarrow \infty} \pi_0(\Psi_t^n x^n) = \pi_0(\Psi_t^\infty(x^\infty))$  because  $\pi_0$  is continuous for pointwise convergence.

Because  $T^{\Gamma_n} \rightarrow T^\infty$  pointwise uniformly, we therefore know that  $\lim_{n \rightarrow \infty} T^{\Gamma_n}(\pi_0(\Psi_t^\infty(x^n))) = T^\infty(\pi_0(\Psi_t^\infty x^\infty))$  pointwise. Thus, by the uniform convergence of the  $\Psi_1^n$  to  $\Psi_1^\infty$ ,

$$\lim_{n \rightarrow \infty} \Psi_t^n x^n = \lim_{n \rightarrow \infty} \Psi_1^n(T^{\Gamma_n}(\pi_0(\Psi_t^n x^n))) = \lim_{n \rightarrow \infty} \Psi_1^\infty(T^{\Gamma_n}(\pi_0(\Psi_t^\infty x^\infty))) = \lim_{n \rightarrow \infty} y^n.$$

We finish the proof of Theorem 9.8 by proving that when each  $\Gamma_n$  contains a minimizer and  $\lim_{n \rightarrow \infty} S_j^n = S_j^\infty$  uniformly in  $C^0(\mathcal{B}_K)$ , then also  $\Gamma_\infty$  contains a minimizer:

*Minimizing property:* Suppose that every ghost circle  $\Gamma_n$  contains a minimizer  $x^n = x^n(\xi_n)$ . This means that for every finite subset  $B \subset \mathbb{Z}^d$  and every  $y: \mathbb{Z}^d \rightarrow \mathbb{R}$  with finite support in  $\dot{B}^{(r)}$  it holds that

$$W_B^n(x^n + y) - W_B^n(x^n) \geq 0, \quad \text{where } W_B^n(x) := \sum_{j \in B} S_j^n(x). \quad (9.14)$$

By compactness of  $\mathcal{B}_K/\mathbb{Z}$ , a subsequence of the  $x^n(\xi_n)$  converges pointwise, say to  $x^\infty = \lim_{j \rightarrow \infty} x^{n_j}(\xi_{n_j})$ . Moreover,  $\lim_{n \rightarrow \infty} W_B^n = W_B^\infty$  uniformly in  $C^0(\mathcal{B}_K)$ . Taking the limit of Eq. (9.14) as  $n_j \rightarrow \infty$  then shows that  $W_B^\infty(x^\infty + y) - W_B^\infty(x^\infty) \geq 0$ . In other words,  $x^\infty$  is a global minimizer.

It remains to prove that  $x^\infty \in \Gamma_\infty$ . This holds because  $x^\infty = \lim_{j \rightarrow \infty} x^{n_j}(\xi_{n_j}) = \lim_{j \rightarrow \infty} \Psi_1^n(T^{\Gamma_n}(\xi_{n_j})) = \Psi_1^\infty(T^\infty(\xi_\infty)) = x^\infty(\xi_\infty)$ , where  $\xi_\infty := \lim_{j \rightarrow \infty} \xi_{n_j}$ .  $\square$

## 10. Gap solutions

In this final section we examine the situation that an Aubry–Mather set  $\mathcal{M}(x) \subset \mathcal{B}_\omega$  has a gap, that is when there are elements  $y^-, y^+ \in \mathcal{M}(x)$  with  $y^- \ll y^+$  such that  $[y^-, y^+]$  does not contain any elements of  $\mathcal{M}(x)$  other than  $y^-$  and  $y^+$ . This situation occurs when  $\omega \in \mathbb{Q}^d$  or when  $\omega \in \mathbb{R}^d \setminus \mathbb{Q}^d$  and  $\mathcal{M}(x)$  is a Cantor set.

The main result of this section is Theorem 10.7 below, which states that either  $[y^-, y^+]$  admits a foliation by global minimizers, or there exists at least one stationary configuration  $z \in [y^-, y^+]$  that is not a global minimizer. This result is more precise than the result of [4], that says that a gap must contain at least one stationary solution. Moreover, the proof below is more geometric, as it makes use of ghost circles.

We start with the following theorem, which is a refinement of a result by Moser [15]. It says that when a gap admits a foliation by stationary points, then all of them are minimizing. Recall that every Aubry–Mather set is contained in a ghost circle.

**Theorem 10.1.** *Let  $[y^-, y^+]$  be a gap in the Aubry–Mather set  $\mathcal{M}(x)$  and let  $\Gamma$  be a ghost circle so that  $\mathcal{M}(x) \subset \Gamma$ . If  $\Gamma^{[y^-, y^+]} := \Gamma \cap [y^-, y^+]$  consists of stationary configurations only, then all of them are global minimizers.*

**Proof.** Assume that  $\Gamma^{[y^-, y^+]}$  consists of stationary points only but that  $w \in \Gamma^{[y^-, y^+]}$  is not a global minimizer. Then there is a finite subset  $B \subset \mathbb{Z}^d$  and a  $z \in \mathbb{R}^{\mathbb{Z}^d}$  with  $\text{supp}(z) \subset \dot{B}^{(r)}$  such that  $W_B(w + z) < W_B(w)$ . Because the function  $z \mapsto W_B(w + z)$  is coercive, it attains its minimum, let's say at a  $Z$  with support in  $\dot{B}^{(r)}$ . By assumption  $Z \neq 0$ . Let's say there is an  $i \in \dot{B}^{(r)}$  for which  $Z_i > 0$ . In the case that  $Z_i < 0$  the proof is similar. We now claim that  $Z$  can be chosen so that  $w + Z \leq y^+$ .

To prove this claim, we remark that when  $m := \min\{w + Z, y^+\}$  and  $M := \max\{w + Z, y^+\}$ , then  $W_B(w + Z) + W_B(y^+) \geq W_B(m) + W_B(M)$ , as in the proof of Lemma 4.4. Because both  $(w + Z) - m$  and  $y^+ - M$  are supported in  $\dot{B}^{(r)}$  and both  $w + Z$  and  $y^+$  minimize  $W_B$  with respect to variations supported in  $\dot{B}^{(r)}$ , it must therefore hold that  $W_B(w + Z) = W_B(m) = W_B(w + \min\{Z, y^+ - w\})$ .

The next step is to define  $y := \inf\{\tilde{y} \in \Gamma \mid \tilde{y} \gg w + Z\}$ . Because  $w, y, y^+ \in \Gamma$ ,  $w + Z \leq y^+$  and  $Z_i > 0$ , it now holds that  $w \ll y \leq y^+$ . At the same time, because  $\Gamma$  is connected,  $y$  touches  $w + Z$ . That is: there is an  $i \in \dot{B}^{(r)}$  so that  $y_i = w_i + Z_i$ , while  $z_k + W_k = z_k < y_k$  for all  $k \notin \dot{B}^{(r)}$ . We claim that this is impossible.

To prove this, choose such an  $i \in \dot{B}^{(r)}$  at which  $y_i = w_i + Z_i$  and a  $k \in \mathbb{Z}^d$  for which  $\|i - k\| = 1$ . Then, because  $y$  is a global stationary point and  $w + Z$  is stationary for  $W_B$  with respect to variations in  $\dot{B}^{(r)}$ , it must be true that

$$\begin{aligned} 0 &= \partial_i W_B(y) - \partial_i W_B(w + Z) \\ &= \sum_{j \in B, l \in \mathbb{Z}^d} \left( \int_0^1 \partial_{i,l} S_j(ty + (1-t)(w + Z)) dt \right) \cdot (y_l - w_l - Z_l) \\ &\geq \int_0^1 \partial_{i,k} S_i(ty + (1-t)(w + Z)) dt \cdot (y_k - w_k - Z_k). \end{aligned}$$

Here, the inequality holds because  $y_i - w_i - Z_i = 0$  and  $\partial_{i,l} S_j \leq 0$  when  $i \neq l$  and  $y_l - w_l - Z_l \geq 0$  for all  $l$ . The twist condition that  $\partial_{i,k} S_i < 0$  then guarantees that  $y_k = w_k + Z_k$ . By induction, one then finds that there is a  $k \notin \dot{B}^{(r)}$  for which  $y_k = w_k + Z_k = w_k$ . This is a contradiction.  $\square$

We will now show that when  $\Gamma^{[y^-, y^+]}$  does not consist of only stationary points, i.e. minimizers, then it contains at least one non-minimizing stationary point. We do this by finding a stationary point of a “renormalized action” function  $W_{[y^-, y^+]} : [y^-, y^+] \rightarrow [0, \infty)$ . In order to define  $W_{[y^-, y^+]}$ , we need the following well-known technical result that states, when applied to  $\mathcal{M} = \mathcal{M}(x)$ , that the gaps of an Aubry–Mather set are uniformly summable:

**Theorem 10.2.** *Let  $\mathcal{M}$  be any strictly ordered, shift-invariant collection of configurations of rotation vector  $\omega \in \mathbb{R}^d$ . Let  $x, y \in \mathcal{M}$  be so that  $x \ll y$  and assume that there exists no  $z \in \mathcal{M}$  with  $x \ll z \ll y$ . Denote  $H_\omega := \{i \in \mathbb{Z}^d \mid \langle \omega, i \rangle \in \mathbb{Z}\}$ . Then*

$$\sum_{i \in \mathbb{Z}^d / H_\omega} |y_i - x_i| \leq 1.$$

**Proof.** We start by remarking that our assumptions on  $\mathcal{M}$  imply that  $x$  and  $y$  are Birkhoff. Now, let  $i$  and  $j$  be representatives of different equivalent classes of  $\mathbb{Z}^d / H_\omega$  and let  $m, n \in \mathbb{Z}$  be arbitrary. Then,  $\langle \omega, j - i \rangle + m - n \neq 0$  and hence, by Proposition 3.8, either  $\tau_{j-i, m-n} y \ll y$  or  $\tau_{j-i, m-n} x \gg x$ . In the first case, actually  $\tau_{j-i, m-n} y \leq x$  because there is no element of  $\mathcal{M}$  between  $x$  and  $y$ . Evaluating the latter inequality at  $i$ , we then obtain  $y_j + m \leq x_i + n$ . In the second case, one finds that  $\tau_{j-i, m-n} x \geq y$ , whence  $x_j + m \geq y_i + n$ . In both cases we find that  $(x_j + m, y_j + m) \cap (x_i + n, y_i + n) = \emptyset$ .

For  $\xi \in \mathbb{R}$ , denote by  $[\xi] := \max\{n \in \mathbb{Z} \mid n \leq \xi\}$  and define  $\bar{x} = \inf_i (x_i - [x_i]) \in [0, 1)$ . Then, clearly  $x_i - [x_i] \geq \bar{x}$ . We claim that  $y_i - [x_i] \leq \bar{x} + 1$ . To prove this, note that  $x_i - [x_i] \leq x_j - [x_j] + 1$  and hence, by the Birkhoff property,  $x \ll \tau_{j-i} x + [x_i] - [x_j] + 1$ . The assumption that  $x$  and  $y$  are consecutive elements of  $\mathcal{M}$  then implies that  $y \leq \tau_{j-i} x + [x_i] - [x_j] + 1$ , that is,  $y_i - [x_i] \leq x_j - [x_j] + 1$ . Hence,  $y_i - [x_i] \leq \bar{x} + 1$ .

This yields, denoting by  $|A|$  the Lebesgue measure of a set  $A \subset \mathbb{R}$ :

$$\begin{aligned} \sum_{i \in \mathbb{Z}^d / H_\omega} |y_i - x_i| &= \sum_{i \in \mathbb{Z}^d / H_\omega} |(x_i - [x_i], y_i - [x_i])| \\ &= \left| \bigcup_{i \in \mathbb{Z}^d / H_\omega} (x_i - [x_i], y_i - [x_i]) \right| \leq |(\bar{x}, \bar{x} + 1)| = 1. \quad \square \end{aligned}$$

For *rationally independent* rotation vectors, for which  $H_\omega = \{0\}$ , Theorem 10.2 was stated for the first time by Moser [14].



For a gap  $[y^-, y^+]$ , with  $y^-, y^+ \in \mathcal{B}_\omega$ , let us define

$$\overline{[y^-, y^+]} := \{y \in [y^-, y^+] \mid \tau_{k,l}y = y \text{ if } \langle \omega, k \rangle + l = 0\}.$$

It is not hard to see that  $\overline{[y^-, y^+]} \subset \mathcal{B}_\omega$ . Namely, when  $y^- < y < y^+$  and  $\langle \omega, k \rangle + l > 0$ , then  $\tau_{k,l}y > \tau_{k,l}y^- \geq y^+ > y$ , where the second inequality holds because  $[y^-, y^+]$  is a gap. Similarly,  $\tau_{k,l}y < y$  when  $\langle \omega, k \rangle + l < 0$ . Hence,  $y$  is Birkhoff once  $\tau_{k,l}y = y$  for all  $k, l$  with  $\langle \omega, k \rangle + l = 0$ . We are now ready to define the renormalized action function:

**Definition 10.3.** When  $[y^-, y^+]$  is a gap, we define  $W_{[y^-, y^+]} : \overline{[y^-, y^+]} \rightarrow [0, \infty)$  by

$$W_{[y^-, y^+]}(y) := \sum_{j \in \mathbb{Z}^d / H_\omega} (S_j(y) - S_j(y^-)).$$

**Proposition 10.4.** For every  $y \in \overline{[y^-, y^+]}$ , the sum  $W_{[y^-, y^+]}(y)$  is absolutely convergent. Moreover,  $W_{[y^-, y^+]}$  is continuous with respect to pointwise convergence.

**Proof.** The compactness of  $\mathcal{B}_\omega$  implies that there is a constant  $D > 0$  so that  $|\partial_k S_j| \leq D$  uniformly on  $\overline{[y^-, y^+]}$ . Thus, we compute that for  $y^1, y^2 \in \overline{[y^-, y^+]}$ ,

$$\begin{aligned} |W_{[y^-, y^+]}(y^1) - W_{[y^-, y^+]}(y^2)| &\leq \sum_{j \in \mathbb{Z}^d / H_\omega} |S_j(y^1) - S_j(y^2)| \\ &\leq \sum_{j \in \mathbb{Z}^d / H_\omega} \sum_{\|k-j\| \leq r} \left| \int_0^1 \partial_k S_j(ty^1 + (1-t)y^2) dt \right| |y_k^1 - y_k^2| \\ &\leq (2r+1)^d D \sum_{k \in \mathbb{Z}^d / H_\omega} |y_k^1 - y_k^2|. \end{aligned}$$

First of all, this implies that  $W_{[y^-, y^+]}(y)$  is well defined and converges absolutely for  $y \in \overline{[y^-, y^+]}$ , because  $W_{[y^-, y^+]}(y^-) = 0$ , by definition, and  $\sum_{k \in \mathbb{Z}^d / H_\omega} |y_k - y_k^-| \leq 1$  by Theorem 10.2.

Secondly, it is now clear that  $W_{[y^-, y^+]}$  is continuous for pointwise convergence, because when  $y^n \in \overline{[y^-, y^+]}$  is a sequence of configurations converging pointwise, say to  $y^\infty$ , then  $\lim_{n \rightarrow \infty} \sum_{j \in \mathbb{Z}^d / H_\omega} |y_j^n - y_j^\infty| = 0$ , as is quite easy to prove, so that  $\lim_{n \rightarrow \infty} W_{[y^-, y^+]}(y^n) = W_{[y^-, y^+]}(y^\infty)$ .  $\square$

The next result is harder to prove:

**Theorem 10.5.** We have  $W_{[y^-, y^+]} \geq 0$ . Moreover, if a configuration  $y \in \overline{[y^-, y^+]}$  is a global minimizer, then  $W_{[y^-, y^+]}(y) = 0$ . In particular,  $W_{[y^-, y^+]}(y^-) = W_{[y^-, y^+]}(y^+) = 0$ .

**Proof.** The proof of this theorem is similar to that of Theorem 4.9. We will sketch it here.

In fact, we will show that when  $y^1, y^2 \in \overline{[y^-, y^+]}$  and  $y^1$  is a global minimizer, then  $W_{[y^-, y^+]}(y^1) \leq W_{[y^-, y^+]}(y^2)$ . Applied to  $y^1 = y^-$ , this shows that  $W_{[y^-, y^+]} \geq 0$ , whereas when applied to  $y^2 = y^-$ , it shows that  $W_{[y^-, y^+]}(y) = 0$  if  $y$  is a global minimizer.

So let  $y^1, y^2 \in \overline{[y^-, y^+]}$  and suppose that  $\varepsilon := W_{[y^-, y^+]}(y^1) - W_{[y^-, y^+]}(y^2) > 0$ . It suffices to show that this implies that  $y^1$  is not a global minimizer. To prove this, let  $(p_1, q_1), \dots, (p_c, q_c)$  be principal

periods for  $\omega$  and let  $B_p \subset \mathbb{Z}^d$  be the fundamental domain for  $\mathbb{Z}^d/H_\omega$  defined by

$$B_p := \{i \in \mathbb{Z}^d \mid 0 \leq \langle i, p_j \rangle < 1 \text{ for all } 1 \leq j \leq c\}.$$

Then  $W_{[y^-, y^+]} = W_{B_p}|_{\overline{[y^-, y^+]}}$  and  $W_{nB_p}(y^1) - W_{nB_p}(y^2) = \varepsilon \cdot n^c$ .

Now define for every  $n > r + 2$  the configuration  $y^n \in [y^-, y^+]$  by

$$y_i^n := \begin{cases} y_i^2 & \text{if } i \in \hat{B}_{np}^{(r)}, \\ y_i^1 & \text{otherwise.} \end{cases}$$

Then  $y^n - y^1$  is clearly supported in the  $r$ -interior of  $B_{np}$  and it is not too hard to show, using the compactness of  $[y^-, y^+]$ , the uniform  $l_1$ -bound on  $[y^-, y^+]$  and the argument of Theorem 4.9, that there is a constant  $E > 0$  so that

$$\begin{aligned} W_{nB_p}(y^1) - W_{nB_p}(y^n) &= (W_{nB_p}(y^1) - W_{nB_p}(y^2)) + (W_{nB_p}(y^2) - W_{nB_p}(y^n)) \\ &> \varepsilon \cdot n^c - E \cdot n^{c-1}. \end{aligned}$$

This means that  $y^1$  is not a global minimizer.  $\square$

One can in fact also prove a variant of Theorem 4.8 that says that if  $W_{[y^-, y^+]}(y) = 0$ , then  $y$  is a global minimizer. Since we do not need this result in this paper, we will not prove it here.

Recall that both  $[y^-, y^+]$  and  $\overline{[y^-, y^+]}$  are invariant under the forward flow of the negative gradient vector field  $-\nabla W$ . But a ghost circle is also invariant under the backward flow. This implies that, if  $\Gamma$  is a ghost circle and  $y^-, y^+ \in \Gamma$  are the endpoints of a gap in an Aubry–Mather set contained in  $\Gamma$ , then  $\Gamma^{[y^-, y^+]}$  is both forward and backward invariant under the negative gradient flow. In order to prove that  $y^-$  and  $y^+$  are not the only fixed points in  $[y^-, y^+]$ , we will now show that  $W_{[y^-, y^+]}$  acts as a Lyapunov function:

**Lemma 10.6.** *Let  $y \in \Gamma^{[y^-, y^+]}$  and denote by  $t \mapsto \Psi_t$  the flow of  $-\nabla W$ . Then  $t \mapsto W_{[y^-, y^+]}(\Psi_t y)$  is continuously differentiable and*

$$\left. \frac{d}{dt} \right|_{t=0} W_{[y^-, y^+]}(\Psi_t y) = - \sum_{i \in \mathbb{Z}^d/H_\omega} (\partial_i W(y))^2.$$

**Proof.** Let us denote, for convenience,  $\|\nabla W(y)\|^2 := \sum_{i \in \mathbb{Z}^d/H_\omega} (\partial_i W(y))^2$ . We will begin by showing that the function  $y \mapsto \|\nabla W(y)\|^2$  is absolutely convergent and continuous on  $\overline{[y^-, y^+]}$ . This is proved by interpolation, as in the proof of Proposition 10.4. More precisely, if  $|\partial_i S_j| \leq D$  and  $|\partial_{i,k} S_j| \leq C$  uniformly on  $\overline{[y^-, y^+]}$ , then  $|\partial_i W| \leq (2r+1)^d D$ , so that

$$\begin{aligned} & \left| \|\nabla W(y^1)\|^2 - \|\nabla W(y^2)\|^2 \right| \\ & \leq \sum_{i \in \mathbb{Z}^d/H_\omega} |\partial_i W(y^1) + \partial_i W(y^2)| \cdot |\partial_i W(y^1) - \partial_i W(y^2)| \\ & \leq \sum_{i \in \mathbb{Z}^d/H_\omega} 2(2r+1)^d D \sum_{\|j-i\| \leq r} \sum_{\|k-j\| \leq r} \left| \int_0^1 \partial_{i,k} S_j(\tau y^1 + (1-\tau)y^2) d\tau \right| \cdot |y_k^1 - y_k^2| \\ & \leq 2CD(2r+1)^{3d} \sum_{k \in \mathbb{Z}^d/H_\omega} |y_k^1 - y_k^2|. \end{aligned}$$

Applied to  $y^1 = y^-$  and  $y^2 = y$  and combined with Theorem 10.2, this implies that  $\|\nabla W(y)\|^2$  is absolutely convergent, because  $\|\nabla W(y^-)\|^2 = 0$ . The continuity for pointwise convergence follows from the argument given in Proposition 10.4. In particular, we now know that  $t \mapsto -\|\nabla W(\psi_t y)\|^2$  is continuous, being the composition of two continuous functions.

The next step is to denote  $y(t) := \psi_t y$ , for  $y \in \Gamma^{[y^-, y^+]}$ , and to observe that

$$S_j(y(t)) - S_j(y) = \int_0^t \frac{d}{d\tau} S_j(y(\tau)) d\tau = - \sum_{\|k-j\| \leq r} \int_0^t \partial_k S_j(y(\tau)) \cdot \partial_k W(y(\tau)) d\tau.$$

Summing this over  $j \in \mathbb{Z}^d/H_\omega$ , we obtain because of the absolute convergence, that

$$W_{[y^-, y^+]}(y(t)) - W_{[y^-, y^+]}(y) = - \int_0^t \|\nabla W(y(\tau))\|^2 d\tau.$$

In particular, because the integrand is continuous, we find that

$$\left. \frac{d}{dt} \right|_{t=0} W_{[y^-, y^+]}(\psi_t y) = \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \|\nabla W(y(\tau))\|^2 d\tau = -\|\nabla W(y)\|^2. \quad \square$$

Proposition 10.4 and Lemma 10.6 combined now lead to the main result of this section:

**Theorem 10.7.** *Let  $[y^-, y^+]$  be a gap in the Aubry–Mather set  $\mathcal{M}(x)$  and let  $\Gamma$  be a ghost circle so that  $\mathcal{M}(x) \subset \Gamma$ . Then either  $\Gamma^{[y^-, y^+]} = \Gamma \cap [y^-, y^+]$  consists of global minimizers only, or there is at least one stationary point  $y \in \Gamma^{[y^-, y^+]}$  that is not a global minimizer.*

**Proof.** By Proposition 10.4, we have that  $W_{[y^-, y^+]} \geq 0$ . When  $W_{[y^-, y^+]}|_{\Gamma^{[y^-, y^+]}} \equiv 0$ , then the flow-invariance of  $\Gamma^{[y^-, y^+]}$  implies that  $W_{[y^-, y^+]}(\psi_t y) = 0$  for all  $y \in \Gamma^{[y^-, y^+]}$  and all  $t \in \mathbb{R}$ . By Proposition 10.6, we then have that  $\|\nabla W(y)\|^2 = 0$ . That is,  $\Gamma^{[y^-, y^+]}$  consists of stationary points only, and hence by Theorem 10.1, it consists of global minimizers only.

Because  $\Gamma^{[y^-, y^+]}$  is compact and  $W_{[y^-, y^+]}$  is continuous, the other possibility is that  $W_{[y^-, y^+]}|_{\Gamma^{[y^-, y^+]}}$  assumes a positive maximum at some point  $y \in \Gamma^{[y^-, y^+]}$  with  $y^- \ll y \ll y^+$ . Proposition 10.4 implies that this  $y$  is not a global minimizer. It is clearly stationary though: if not, then  $\|\nabla W(y)\|^2 > 0$ , so that by continuity of  $t \mapsto \|\nabla W(\psi_t y)\|^2$ , we have for each  $t < 0$  that

$$W_{[y^-, y^+]}(\psi_t y) - W_{[y^-, y^+]}(y) = - \int_0^t \|\nabla W(\psi_\tau y)\|^2 d\tau > 0.$$

Because  $\psi_t y \in \Gamma^{[y^-, y^+]}$ , this contradicts that  $y$  is a maximizer of  $W_{[y^-, y^+]}|_{\Gamma^{[y^-, y^+]}}$ .  $\square$

At this moment it is unclear to us whether a gap in an Aubry–Mather set can be foliated by stationary points – which therefore all have to be nonrecurrent global minimizers.

## Acknowledgments

We would like to thank our colleagues of the Department of Mathematics at VU University Amsterdam for their continuous support. A large part of this paper was written during the authors' visit to the Department of Mathematics and Statistics at Boston University. This research was funded by the Dutch Science Foundation NWO.

## Appendix A. Twist maps

Variational monotone recurrence relations do not only arise in statistical mechanics or as discretized PDEs: the case of dimension  $d = 1$  is relevant for the theory of twist maps of the cylinder. The latter arise for instance in the study of convex billiards and as Poincaré maps of Hamiltonian systems. In this short descriptive appendix, we will briefly review these topics. The informed reader can skip this appendix and we refer to [13] or [9] for more detailed proofs of our statements, as well as for a more comprehensive introduction to the topic.

Let us denote by  $\mathbb{A} := \mathbb{R}/\mathbb{Z} \times \mathbb{R}$  the standard cylinder, with coordinates  $(x \bmod 1, y)$  and bundle projection  $\pi : \mathbb{A} \rightarrow \mathbb{R}/\mathbb{Z}$  given by  $(x \bmod 1, y) \mapsto x \bmod 1$ . The lift  $\tilde{\pi} : \mathbb{R}^2 \rightarrow \mathbb{R}$  of  $\pi$  to the universal covering spaces sends  $(x, y)$  to  $x$ .

Recall that a smooth cylinder map  $T : \mathbb{A} \rightarrow \mathbb{A}$  allows for a lift  $\tilde{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , with the property that  $\tilde{T}(x, y) \bmod(1, 0) = T(x \bmod 1, y)$ . This implies that  $\tilde{T}(x + 1, y) = \tilde{T}(x, y) + (n, 0)$ , where  $n$  is the degree of  $T$ , and moreover that  $\tilde{T}$  is unique modulo constants of the form  $(m, 0)$  with  $m \in \mathbb{Z}$ .

**Definition A.1.** We call a cylinder map  $T : \mathbb{A} \rightarrow \mathbb{A}$  an *exact symplectic positive twist map* if it satisfies conditions 1, 2 and 3 below.

1. *Degree one:*  $\tilde{T}(x + 1, y) = \tilde{T}(x, y) + (1, 0)$ .

Condition 1 is true if and only if  $T$  is homotopic to the identity map of  $\mathbb{A}$ .

2. *Exact symplectic:* The one-form  $T^*(y dx) - y dx$  is exact.

Condition 2 implies that there is a so-called *generating function*, denoted  $s : \mathbb{A} \rightarrow \mathbb{R}$ , such that  $T^*(y dx) - y dx = ds$ . We will denote its lift by  $\tilde{s} : \mathbb{R}^2 \rightarrow \mathbb{R}$ . This lift satisfies the identity  $\tilde{s}(x, y) = s(x \bmod 1, y)$ . In particular,  $\tilde{s}(x + 1, y) = \tilde{s}(x, y)$ .

Geometrically, condition 2 can be interpreted as follows: one can show that when condition 1 holds, then condition 2 is true if and only if  $T$  preserves the volume form  $dy \wedge dx$  and moreover has the property that the volume enclosed by the cycles  $\gamma(t) := (t \bmod 1, 0)$  and its homotopic image  $T \circ \gamma$  is equal to zero.

Moreover, conditions 1 and 2 hold if and only if  $T$  is a so-called Hamiltonian map, i.e.  $T$  is the time-1 flow of a time-1-periodic Hamiltonian vector field  $X_{H(t)}$  on  $\mathbb{A}$ .

To formulate the last condition, let us call  $\tilde{T}(x, y) = (X(x, y), Y(x, y))$ .

3. *Positive twist:* The map  $(x, y) \mapsto (x, X(x, y)) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a diffeomorphism. This implies that  $\partial_y X \neq 0$ . We require that  $\partial_y X > 0$ .

Condition 3 says that  $T$  twists each fiber  $\{x \bmod 1\} \times \mathbb{R} \subset \mathbb{A}$  around the cylinder  $\mathbb{A}$  “in the positive direction”. We will denote the inverse of the map  $(x, y) \mapsto (x, X(x, y))$  by  $(x, X) \mapsto (x, y(x, X))$ .

In fact, condition 3 allows us to define the function  $S : \mathbb{R}^2 \rightarrow \mathbb{R}$  by  $S(x, X) := \tilde{s}(x, y(x, X))$ . The function  $S$  is called the *generating function* of the twist map  $T$ .

The following well-known theorem is crucial in the theory of twist maps. It states that, in order to find orbits of exact symplectic positive twist maps of the cylinder, one needs to solve a variational monotone recurrence relation in dimension  $d = 1$ . For completeness, we have included a brief proof of this statement.

**Theorem A.2.** The sequence  $i \mapsto (x_i, y_i) \in \mathbb{R}^2$  is an orbit of  $\tilde{T}$  if and only if

- i) For all  $M, N \in \mathbb{Z}$  with  $M < N$  the sequence  $i \mapsto x_i$  is a stationary point of the finite action

$$W_{M,N}(x) := \sum_{i=M+1}^N S_i(x), \quad \text{with } S_i(x) := S(x_{i-1}, x_i),$$

for variations of  $x$  with fixed endpoints  $x_M$  and  $x_N$ .

ii) It holds that  $y_i = -\partial_X S(x_i, x_{i+1})$  for all  $i$ .

Moreover, one has that  $S(x+1, X+1) = S(x, X)$  and  $\partial_{x,X} S < 0$ .

**Proof.** Recall the notation  $\tilde{T}(x, y) = (X(x, y), Y(x, y))$ , the diffeomorphism  $(x, y) \mapsto (x, X(x, y))$  with its inverse  $(x, X) \mapsto (x, y(x, X))$  and the definition  $S(x, X) = \tilde{s}(x, y(x, X))$ . Then the equality  $T^*(y dx) = y dx + dS$  in the coordinate system  $(x, y)$  becomes  $Y dX = y dx + dS$  in the coordinates  $(x, X)$ , viewing  $y = y(x, X)$  and  $Y = Y(x, y(x, X))$  as functions of  $x$  and  $X$ . Writing  $dS = \partial_x S dx + \partial_X S dX$ , we thus obtain that

$$Y = \partial_X S \quad \text{and} \quad y = -\partial_x S. \quad (\text{A.1})$$

Now let  $i \mapsto (x_i, y_i)$  be an arbitrary sequence and define  $X_i := X(x_{i-1}, y_{i-1})$  and  $Y_i := Y(x_{i-1}, y_{i-1})$ . Then  $(x_i, y_i)$  is an orbit of  $\tilde{T}$  if and only if  $x_i = X_i$  and  $y_i = Y_i$  for all  $i$ . According to formula (A.1), we have that  $Y_i = Y(x_{i-1}, y_{i-1}) = \partial_X S(x_{i-1}, X(x_{i-1}, y_{i-1})) = \partial_X S(x_{i-1}, X_i)$  and  $y_i = -\partial_x S(x_i, X(x_i, y_i)) = -\partial_x S(x_i, X_{i+1})$ . Thus,  $x_i = X_i$  and  $y_i = Y_i$  if and only if  $y_i = -\partial_x S(x_i, x_{i+1})$  and

$$\partial_X S(x_{i-1}, x_i) + \partial_x S(x_i, x_{i+1}) = 0 \quad \text{for all } i \in \mathbb{Z}. \quad (\text{A.2})$$

Formula (A.2) holds if and only if  $i \mapsto x_i$  is stationary for all the  $W_{N,M}$  defined above.

Because  $T$  has degree one, we have that  $X(x+1, y) = X(x, y) + 1$ . The function  $y(x, X)$  is defined implicitly by the relation  $X(x, y(x, X)) = X$ , and therefore we see that  $y(x+1, X+1) = y(x, X)$ . Thus, the generating function satisfies  $S(x+1, X+1) = \tilde{s}(x+1, y(x+1, X+1)) = \tilde{s}(x, y(x, X)) = S(x, X)$ .

Finally, formula (A.1) implies that  $\partial_{x,X} S = -\partial_X y = -(\partial_y X)^{-1} < 0$ .  $\square$

### A.1. Examples of twist maps

Perhaps the most famous example of an exact symplectic twist map is the *Chirikov standard map*. Given a 1-periodic function  $V = V(x)$ , it is defined as  $T_V : \mathbb{A} \rightarrow \mathbb{A}$  by

$$T_V(x, y) = (x + y + 2V'(x) \bmod 1, y + 2V'(x)).$$

It turns out that its generating function is  $S(x, X) := \frac{1}{2}(x - X)^2 + 2V(x)$ . In other words, the variational monotone recurrence relation corresponding to  $T_V$  is exactly the Frenkel–Kontorova equation in dimension  $d = 1$ , given by  $S_i(x) = \frac{1}{4}(x_{i-1} - x_i)^2 + V(x_i)$ . By the way, the “standard” is to choose  $V(x) = \frac{k}{8\pi^2} \cos(2\pi x)$ , for some parameter  $k \geq 0$ . This produces the map  $(x \bmod 1, y) \mapsto (x + y - \frac{k}{2\pi} \sin(2\pi x) \bmod 1, y - \frac{k}{2\pi} \sin(2\pi x))$ .

Another application of the theory of twist maps arises in the context of *convex billiards*, cf. [16]. The configuration space of such a billiard consists of the arclength parameters  $x \in \mathbb{R}/\mathbb{Z}$  that describe the position of the billiard ball along the boundary of the billiard at the moment of reflection and angles  $y \in (0, \pi)$  measuring the direction of the outgoing billiard trajectory with respect to the tangent line to the billiard at  $x$ . Then the motion of a billiard ball is described by an exact symplectic positive twist map  $T : (x_i, y_i) \mapsto (x_{i+1}, y_{i+1})$ . The variational structure of this problem follows as the rule “angle of incidence = angle of reflection” is derived from the variational principle that a billiard ball travels along “shortest paths”. The positive twist condition  $\frac{\partial x_{k+1}}{\partial y_k} > 0$  should be obvious from Fig. 1.

Finally, under generic conditions, the Poincaré return map of a degree two of freedom Hamiltonian system near an elliptic equilibrium point is an exact symplectic twist map. In this case, the corresponding twist map is actually close to *integrable*, so that it allows for the application of various kinds of perturbation theory. Again, we refer to [13] for more details.

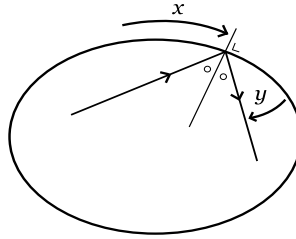


Fig. 1. An ellipse-shaped convex billiard and part of a billiard trajectory.

## References

- [1] Sigurd B. Angenent, The periodic orbits of an area preserving twist map, *Comm. Math. Phys.* 115 (3) (1988) 353–374.
- [2] Victor Bangert, A uniqueness theorem for  $\mathbb{Z}^n$  periodic variational problems, *Comment. Math. Helv.* 62 (4) (1987) 511–531.
- [3] Mikhail L'vovich Blank, Metric properties of minimal solutions of discrete periodical variational problems, *Nonlinearity* 2 (1) (1989) 1–22.
- [4] Rafael de la Llave, Enrico Valdinoci, Critical points inside the gaps of ground state laminations in statistical mechanics, *J. Stat. Phys.* 129 (1) (2007) 81–119.
- [5] Rafael de la Llave, Enrico Valdinoci, Ground states and critical points for Aubry–Mather theory in statistical mechanics, preprint, 2007.
- [6] Lawrence C. Evans, *Partial Differential Equations*, American Mathematical Society, 1998.
- [7] Christophe Golé, A new proof of the Aubry–Mather's theorem, *Math. Z.* 210 (1991) 441–448.
- [8] Christophe Golé, Ghost circles for twist maps, *J. Differential Equations* 97 (1) (1992) 140–173.
- [9] Christophe Golé, *Symplectic Twist Maps*, Transl. Math. Monogr., World Scientific Publishing Co. Pte. Ltd., 2001.
- [10] Morris W. Hirsch, *Differential Topology*, Springer, 1976.
- [11] Hans Koch, Rafael de la Llave, Charles Radin, Aubry–Mather theory for functions on lattices, preprint.
- [12] R.S. MacKay, I.C. Percival, Converse KAM: Theory and practice, *Comm. Math. Phys.* 98 (4) (1985) 469–512.
- [13] John N. Mather, Giovanni Forni, Action Minimizing Orbits in Hamiltonian Systems, *Lecture Notes in Math.*, 1994, pp. 92–186.
- [14] Jürgen Moser, Minimal solutions of variational problems on a torus, *Ann. Inst. Henri Poincaré* 3 (3) (1986) 229–272.
- [15] Jürgen Moser, *Minimal Foliations on a Torus*, vol. 1365/1989, Springer, Berlin/Heidelberg, 1989, pp. 62–99.
- [16] Serge Tabachnikov, *Geometry and Billiards*, Stud. Math. Libr., vol. 30, American Mathematical Society, Providence, RI, 2005.